Mapping Networks to Probability Distributions in the Economy

A dissertation presented

by

Janelle Dana Schlossberger

to

The Committee for the PhD in Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2019

Dissertation Advisor:
Professor Tomasz Strzalecki

Author:
Janelle Dana Schlossberger

**Mapping Networks to
Probability Distributions in the Economy**

Abstract

This dissertation develops and applies a set of theoretical tools that allows us to explicitly map the topologies of networks in the economy to different probability distributions of interest. The first chapter, "The Distribution of Outcomes for a Networked Economy," develops a set of tools for mapping the topology of a network to a probability distribution of possible outcomes for the economy. I adapt these tools to study locally formed macroeconomic sentiment and how agents' interaction structure shapes the capacity for there to exist non-fundamental swings in aggregate macroeconomic sentiment, with implications for our understanding of animal spirits. I can apply these tools to analyze complex systems in closed form and to construct error bounds about the paths of aggregated networked economies. In the second chapter, "The Distribution of Multipliers in a Networked Economy," there is a policymaking actor who wants to increase the aggregate action in a networked population of N agents. To achieve that goal, the policymaker implements a policy targeting $n < N$ agents. This second chapter studies how the topology of agents' interaction network shapes the distributions of possible policy-induced aggregate actions and economic multipliers. I study a general networked setting and three environments with network-based interaction: (1) strategic complements and substitutes, (2) coordination and anti-coordination, and (3) production. Given n, for

each environment, I map the network topology to distributions of possible resulting aggregate actions and multipliers. The third chapter, "Comprehensively Stress Testing the Economy," addresses two main weaknesses in the Federal Reserve's stress testing approach: (1) the number of stress tests faced by each financial institution is quite small, and (2) the Federal Reserve's toolkit is not sufficiently macroprudential. Employing a macroprudential approach, this chapter shows how to massively increase the total number of stress tests without increasing the computational burden. I generate classes of stress tests with large cardinalities; for each class, I construct probability distributions that capture the full range of possible balance sheet effects for individual financial institutions and the overall financial system. This approach shows how the topologies of bipartite networks linking financial institutions to assets shape stress tests' effects.

Contents

List of Figures

Acknowledgments

I would like to thank my dissertation committee members, David Laibson, Xavier Gabaix, Tomasz Strzalecki, Emmanuel Farhi, and Matthew Rabin, for their feedback and guidance.

To my mother Lois,
my father Gary, and
my brother Matthew,
for their love and support

Introduction

In my dissertation, I develop and apply a set of theoretical tools that allows us to explicitly map the topologies of networks in the economy to different probability distributions of interest. Below, I discuss the three chapters of my dissertation in greater detail.

In the first chapter, I develop a set of mathematical tools that allows us to map the topology of an economic network to a probability distribution of possible outcomes for the economy. To generate this mapping from network topology to probability distribution, I focus on a class of economies that has the following three features: (1) a population of N agents, each with a binary-valued attribute, (2) a network on which these N agents are organized, and (3) decision-making by each networked agent that depends on the local relative frequency of the attribute. I begin by constructing in closed form the distribution of possible local relative frequencies of the attribute given the topology of the underlying network and the attribute's global relative frequency. The topology of the underlying network determines the extent to which the local relative frequency of the attribute can deviate from its global relative frequency, thereby determining the extent to which the outcome of the economy can deviate from a benchmark outcome. Then, given this distribution and agents' decision-making behavior, I construct the distribution of possible outcomes for the economy. For realistic agent interaction structures featuring a very large population of agents, the distribution of outcomes is meaningfully non-degenerate. I adapt the theoretical framework and mathematical tools developed in this work to study locally formed macroeconomic sentiment and how agents' interaction structure shapes the capacity for there to exist non-fundamental swings in aggregate macroeconomic sentiment, with implications for our understanding of animal spirits. I can moreover apply these tools to analyze complex economic systems in closed form and to construct error bounds about the paths of aggregated networked economies.

In the second chapter, I study how the topology of agents' interaction network shapes the distribution of possible policy-induced aggregate actions and the distribution of possible policy-induced economic multipliers. There is a policymaking actor who wants to increase the aggregate action in a networked population of N agents. To achieve that goal, the policymaker implements a policy targeting $n < N$ agents. Since the agents in the population are networked and the actions that they take are interdependent, the aggregate action and the policy-specific economic multiplier crucially depend on which group of n agents gets targeted by the policy. In this work, I therefore study how the topology of agents' interaction network shapes the corresponding probability distributions of possible aggregate actions and economic multipliers for any given policy. I study a general networked setting and three broad environments with network-based interaction: (1) strategic complements and substitutes, (2) coordination and anti-coordination, and (3) production. I show that the mathematics is the same across all three environments; the general networked setting nests each of these three environments. Given n, for each environment, I map the topology of agents' interaction network to the distribution of possible resulting aggregate actions and economic multipliers. The statistical features of these distributions provide crucial information about a policy's efficacy, and I can compute these features in closed form. I can also rank networks so that the outside actor's policy is more effective the higher ranked the network. I show how non-trivial network topologies generate negative multipliers. Across all three environments, there is often a non-zero probability that the enacted policy will reduce the aggregate action below its no-intervention level, and I can compute this probability in closed form.

In the third chapter, I study how the topologies of bipartite networks linking financial institutions to assets shape stress tests' effects on the financial system. In response to the global financial crisis of 2008, the Federal Reserve decided to develop and implement stress tests to assess the soundness of the financial system. Each stress test involves crafting a potential real-world scenario and then quantifying the scenario's effect on both financial actors in the economy and the financial system as a whole. There currently exist two weaknesses in the Federal Reserve's stress testing approach. First, the number of stress tests faced by each financial institution is quite small, with many such stress test scenarios mimicking past historical events that are not necessarily reflective of future situations. Second, the Federal Reserve's toolkit is not sufficiently macroprudential in nature, even though the financial crisis did cause many central banks to nominally transition from a microprudential regulatory approach to a macroprudential regulatory approach. In this third chapter, I tackle these two issues. I show how to massively increase the number and types of possible stress tests without increasing the computational burden. To do this, I generate classes of stress tests with potentially very large cardinalities. For each class of stress tests, I then construct in closed form probability distributions that capture the range of possible balance sheet effects both for each individual financial institution and for the entire financial system. The approach that I take towards increasing the number of stress tests is fundamentally macroprudential.

Chapter 1

The Distribution of Outcomes for a Networked Economy

1.1 Introduction

In this chapter, we study an economy that has a population of agents organized on a network. Given the features of the economy's agents, the topology of the underlying network, and agents' decision-making behavior, we are interested in the distribution of possible outcomes for the economy. The existence of a network structure introduces complications into the economic system; the outcome of the economy is inherently dependent on the topology of the underlying network, but it is not immediately apparent how the topology shapes the distribution of outcomes. This work provides that mapping from network structure to probability distribution.

The main contribution of this work is that it introduces mathematics that enables us to derive in closed form the probability distribution of possible outcomes for the economy from the topology of agents' interaction network. The high-level innovation of this work is that it develops a set of tools that mathematically links two fields:

networks to statistics, or more precisely, networks to probability distributions. The topology of the network directly affects the shape of the probability distribution of outcomes, and this work makes that relationship explicit. We can carry out this mapping from network to probability distribution for all feasible network topologies. The economy in which the network is embedded also affects the mapping, and we take the features of the economy into account.

To develop this mapping, we focus on a class of economic systems that has three distinguishing characteristics. First, the economy has a population of N agents, each of whom has a binary-valued attribute. Second, these N agents are organized on a network. Third, each networked agent's decision-making depends on the local relative frequency of the attribute. Since agents have different positions on the network, they potentially have different local relative frequencies of the attribute arising from different network neighborhoods, and this can lead the agents to make different decisions. Each economy in this class also has an aggregate feature. That aggregate feature is the global relative frequency of the attribute. There are $n \leq N$ agents with the attribute's unit value, so the global relative frequency of the attribute is $f = \frac{n}{N}$. The exogenous objects in this work are the population size, the underlying network structure, the global relative frequency of the attribute, and agents' decision-making behavior, while our endogenous object of interest is the probability distribution of possible outcomes for the economy.

In this work, we are mapping the topology of agents' interaction network to a distribution of outcomes for the economy. In general, we have a non-degenerate probability distribution because there is more than one possible outcome for the economy. This multiplicity of outcomes arises because there are combinatorially many possible configurations, or arrangements, of the binary-valued attribute among agents consistent with the attribute's global relative frequency. When

$f = \frac{n}{N}$, there are $\binom{N}{n}$ possible configurations. We can imagine that the outcome of the economy changes with the particular configuration of the attribute. As the configuration changes, a different subset of agents has the attribute's unit value, which generates a potential adjustment to the local relative frequency of the attribute for each agent. Agents choose actions based on that local relative frequency of the attribute, so as the configuration changes, we potentially have a shift in agents' actions, which leads to a different outcome for the economy.

The two objects that we focus on in this work are the local relative frequency of the attribute and the outcome for the economy. Holding fixed the attribute's global relative frequency, there are combinatorially many possible configurations and for each configuration, there is an associated local relative frequency of the attribute. Therefore, given the attribute's global relative frequency and the structure of the underlying network, we can construct an entire probability distribution of possible local relative frequencies of the attribute. We refer to the distribution of possible local relative frequencies of the attribute as a *precursor distribution* because its construction precedes our construction of the distribution of possible outcomes for the economy. Once we have computed the precursor distribution, we can then construct the distribution of possible outcomes for the economy given agents' decision-making behavior. Our precursor distribution characterizes the extent to which the attribute's local relative frequency deviates in either direction away from its global relative frequency. The capacity for such variation in the attribute's local relative frequency depends on the underlying network structure, and it determines the extent to which the outcome of the economy can deviate from a benchmark outcome.[1] If we have sufficient variation in the local relative frequency of the attribute, then there is an

[1]The benchmark outcome is the one that results if we ignore the underlying configuration and only take into account the aggregate properties of the system, namely the attribute's global relative frequency.

entire non-degenerate distribution of possible outcomes for the economy.

We are interested in characterizing the distribution of possible outcomes for the economy, but we have not yet made explicit what an outcome is exactly. The outcome for an economy is situational. For example, it might be the aggregate action taken by all agents in the population, or it might instead be the action of a single agent of interest. Alternatively, the outcome of the economy might follow from the outcome of an event; for instance, it might follow from the outcome of an economy-wide political election with two possible candidates. In such a setting, there would be two possible outcomes for the economy, and the probability that each outcome occurs is equal to the probability that the corresponding candidate wins the election.

The main technical contributions of this work are as follows. We characterize the shape and properties, including the higher-order statistical features, of our precursor distribution for every feasible network structure, population size, and global prevalence of the binary-valued attribute. We determine those network topologies for which the local relative frequency of the attribute is invariant to configuration, which makes the precursor distribution degenerate. More generally, we statistically characterize the precursor distribution when every configuration is equally likely and when every configuration occurs with some arbitrary probability. Once we have characterized this distribution of possible local relative frequencies of the attribute in full generality, we then study the distribution of possible outcomes for the economy. For certain classes of agent actions, we can provide a closed-form representation of this distribution of possible outcomes. We can characterize this distribution of outcomes for all feasible network structures, population sizes, and global prevalences of the attribute in the population. To the extent that there is variation in the local relative frequency of the attribute across configurations, there can then be significant variation in the economy's outcome, holding f fixed.

When our probability distribution of possible outcomes is non-degenerate, meaning that the outcome of the economy varies with configuration, we consider the economic system to be configuration dependent. Configuration dependence enables the existence of phenomena that would otherwise not occur if we only considered the system's aggregate features. It adds richness to our models of the economy because there is an entire distribution of possible outcomes consistent with our uniquely valued aggregate feature. If we ignored the inherent configuration dependence of the economy, then there would only be one possible outcome.

The tools that we develop in this work enable us to form insights into: (1) complex economic systems and (2) aggregated economies. First, the mathematical machinery allows us to unpack complex systems, and more specifically, complex economic systems featuring network-based agent interaction. Economies with agent-based interaction are quite complicated, as there are myriad ways that these systems can possibly evolve. The tools developed in this work enable the closed-form analysis of such complex economic systems. We can collapse the complexities of agent-based interaction into a simple probability distribution that characterizes how the system will evolve.

Second, these tools allow us to assess aggregate treatments of economic systems and quantify their incompleteness. The class of models that we study in the present work has features that enable direct comparison with aggregate models of the economy. Here, we have a population of N agents whose decision-making behavior can be aggregated, and there is an aggregate feature, f, built up from the attributes of the underlying set of agents. Aggregate models of the economy similarly feature a population of N agents whose actions can be aggregated, and the corresponding representative agent makes decisions based on the aggregated characteristics of the system. In such aggregate models of the economy, the action taken by the representative agent given the system's aggregate characteristics is unique; there is a single outcome. However, for the class of models in the present work, even though the economy has a parallel structure, there is an entire non-degenerate distribution of possible outcomes for the economy that is consistent with the system's uniquely valued aggregate feature, f.

This work shows how aggregate treatments of economic systems can lead to characterizations that are, in general, incomplete. Rather than the aggregate economic system having a unique outcome determined by the system's aggregate features, there is instead an entire distribution of possible outcomes centered about that original benchmark outcome. Using the tools developed in this work, we can introduce a configurational error bound and place that error bound about the benchmark outcome of the aggregate economy to account for the multiplicity of possible outcomes. In particular, we construct this error bound for aggregated systems with networked agents who make local decisions. The size of this error bound depends on the underlying network's topology. By incorporating this error bound, we allow for a more complete and a more nuanced understanding of the phenomena that aggregate models of the economy seek to study. As a result, two systems with the same aggregate features can evolve differently due to differences in their underlying configurations; the configurational error bound that we construct accounts for this variation in the two economic outcomes relative to each other and relative to the benchmark outcome.

We use these theoretical findings to study locally formed macroeconomic sentiments, election outcomes, and animal spirits. In our applied setting, the binary-valued attribute denotes employment status, and each agent makes a voting decision that depends on his or her local unemployment rate. This local unemployment rate is a proxy for individual macroeconomic sentiment and the average local unemployment rate is a proxy for aggregate macroeconomic sentiment. In our model, the fundamentals of the economy, namely the economy's global unemployment rate, alone favor the election of one candidate with certainty. However, if the overall level of macroeconomic sentiment in the economy sufficiently varies with the underlying configuration of unemployment, then there can be more than one possible election outcome and more than one possible outcome for the economy. In a setting with 137.5 million agents, which is the number of voters in the 2016 U.S. presidential election, we find that the distribution of average local unemployment rates is strongly non-degenerate. The variation in this average local unemployment rate is sufficiently large that it can actually mimic variations in business cycle conditions. As a result, the election outcome depends on the particular configuration of unemployment. Such non-degeneracy of the distribution of outcomes for very large N emerges from both high variance of in-degrees and heavy-tailedness of weighted in-degrees for the calibrated social observation network.

By assuming that agents form macroeconomic sentiment from their local unemployment rates, we are able to quantify the extent to which aggregate sentiment is positive or negative given the economy's fundamentals; aggregate sentiment is positive, that is, there are waves of optimism, when the average local unemployment rate is less than the global unemployment rate, while aggregate sentiment is negative, that is, there are waves of pessimism, when the average local unemployment rate is greater than the global unemployment rate. We can quantify this deviation in the average local unemployment rate from the actual unemployment rate and therefore quantify deviations in aggregate sentiment away from a level that is commensurate with the economy's fundamentals. We show how the underlying interaction structure among agents in the economy shapes the capacity for there to exist these non-fundamental swings in aggregate sentiment for all population sizes. We thus offer a mechanism for the formation of individual and aggregate macroeconomic sentiment that essentially microfounds animal spirits.

1.1.1 Relation to the Literature

This work interfaces with four different strands of the literature: (1) complex economic systems, (2) networks, (3) aggregation, and (4) macroeconomic sentiment. Research in the area of complex economic systems includes Granovetter (1978), Brock and Durlauf (2001a), Bisin et al. (2004), and Horst and Scheinkman (2004). These works all feature some form of agent interaction, namely agents choosing actions that depend on the actions of other agents. These works take great care to establish the existence of equilibria in such settings. For these works, the equilibrium outcome is the object of study. The present work meanwhile has a different focus; the object of interest in the present work is the distribution of possible outcomes for the system.

This work contributes to research on networks by developing a set of tools that allows us to mathematically link the field of networks to the field of statistics, and in particular, the area of probability distributions. This work is innovative in that regard. We can also relate the present work to recent research on network-based social learning. Recent papers in that area include Gale and Kariv (2003), Golub and Jackson (2010), Acemoglu, Ozdaglar, and ParandehGheibi (2010), Acemoglu, Dahleh, Lobel, and Ozdaglar (2011), Banerjee, Breza, Chandrasekhar, and Mobius (2016), Harel et al. (2017), and Chandrasekhar, Larreguy, and Xandri (2018). The present work provides theoretical results that enhance our understanding of DeGroot learning. Through the tools developed in the present work, we are able to construct the entire distribution of possible consensus learned values, including the higher-order features of this distribution, for any population size, and we can determine how the topology of the network shapes the capacity for there to be learning and mis-learning.

Research in the area of aggregation tends to examine whether aggregate fluctuations in output can arise from micro-level shocks, or if an aggregate parameter is needed in models of the macroeconomy to generate sufficiently sizable aggregate fluctuations. Papers include Bak, Chen, Scheinkman, and Woodford (1993), Scheinkman and Woodford (1994), Horvath (1998, 2000), Gabaix (2011), and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012). These papers offer different mechanisms that slow the rate at which the law of large numbers applies. They study the role of sectoral production networks, granularity in firm size, and non-convexities in production technologies coupled with local interaction among sectors to generate adequate fluctuations in output that persist even as the economy becomes increasingly disaggregated. Recently, research in this area has also been trying to study how microeconomic shocks shape the higher-order features of the output distribution; recent work includes Acemoglu, Ozdaglar, and Tahbaz-Salehi (2017) and Baqaee and Farhi (2018). The present work tackles this issue of aggregation as well. It examines the extent to which the distribution of possible outcomes for the economy remains non-degenerate in a large-N setting. We develop theoretical results that characterize the variance and the CDF of this distribution, including its higher-order features, for every possible population size and network topology, including the limit as $N \to \infty$. We can explicitly examine how the topological features of the interaction network shape the capacity for the distribution of outcomes for the economy to remain approximately non-degenerate, even for large N.

Recent research in the area of macroeconomic sentiment includes Barsky and Sims (2012), Angeletos and La'O (2013), Benhabib, Wang, and Wen (2015), Huo and Takayama (2015), Acharya, Benhabib, and Huo (2017), Angeletos, Collard, and Dellas (2017), and Milani (2017). In the theoretical domain, these works define mathematically what it means for economies to have sentiment, consumer confidence, and/or animal spirits. In the empirical domain, the literature has tried to determine the extent to which realistically calibrated shocks to sentiment and/or consumer confidence can impact and generate reasonable fluctuations in macroeconomic aggregates. The present work interacts with the existing literature by providing a simple mechanism for the formation of macroeconomic sentiment among agents. This mechanism allows us to quantify the extent to which individual sentiment and aggregate sentiment deviates from a level that is commensurate with economic fundamentals. The present work also shows how the underlying interaction structure among agents shapes the capacity for there to exist non-fundamental swings in aggregate sentiment.

1.1.2 Outline of Chapter

Section 1.2 provides notation and definitions, introduces the class of problems that we later mathematically solve, and works through two illustrative examples. Section 1.3 applies this class of problems towards understanding macroeconomic sentiments and political election outcomes. It studies how there can be sizable configuration-induced variations in macroeconomic sentiment in a large-N economy for fixed economic fundamentals and the resulting impact on election outcomes. After exploring this application, Section 1.4 begins to develop the mathematics that enables us to solve our class of problems. Section 1.5 first characterizes the null setting in which the particular configuration of the attribute among agents is irrelevant. It identifies those conditions for which the distribution of possible local relative frequencies of the attribute is either degenerate or invariant to configuration. Sections 1.6 and 1.7 then present the tools that enable us to characterize the distribution of possible local relative frequencies of the attribute when it is non-degenerate and every configuration is either equally or not equally likely to occur. Section 1.8 studies the distribution of possible outcomes for the economy given agents' decision-making behavior, and Section 1.9 concludes.

1.2 Model

We begin by introducing the notation and definitions that will be used throughout this paper. We then proceed to develop our guiding theoretical framework, highlighting the objects of interest that emerge. We conclude this section by working through two examples that make the theoretical framework and the objects of interest even more precise in an applied setting.

1.2.1 Notation and Definitions

The cardinality of a set $\mathcal{X}$ is $|\mathcal{X}|$. A *multiset* is an object similar to a set, but it allows for multiple instances of each of its elements. Vector $\mathbf{x}$ is a column vector by default. The i^{th} element of vector $\mathbf{x}$ is x_i or $[\mathbf{x}]_i$. The ij^{th} element of matrix $\mathbf{X}$ is $[\mathbf{X}]_{ij}$, the i^{th} row of $\mathbf{X}$ is $[\mathbf{X}]_{i*}$ and the j^{th} column of $\mathbf{X}$ is $[\mathbf{X}]_{*j}$. The identity matrix is $\mathbf{I}$, the column vector whose elements all equal 1 is $\mathbf{1}$, and the unit vector $\mathbf{e}_i$ has $[\mathbf{e}_i]_j = 1$ for $i = j$ and $[\mathbf{e}_i]_j = 0$ otherwise. Matrix $\mathbf{X}$ is *row-stochastic* if $\mathbf{X1} = \mathbf{1}$ and all matrix elements of $\mathbf{X}$ are non-negative. Matrix $\mathbf{X}$ is *doubly stochastic* if it is both row-stochastic and column-stochastic, that is, $\mathbf{X1} = \mathbf{1}$, $\mathbf{X}^T\mathbf{1} = \mathbf{1}$, and all matrix elements of $\mathbf{X}$ are non-negative. Non-negative matrix $\mathbf{X}$ is *primitive* if there exists an integer $q \geq 1$ such that $[\mathbf{X}^q]_{ij} > 0$ for all matrix elements in $\mathbf{X}^q$. $x(N) \sim y(N)$ *w.h.p.* (that is, $x(N)$ is asymptotically equivalent to $y(N)$ with high probability) if $\Pr\left(\frac{x(N)}{y(N)} \to 1\right) \to 1$ as

$N \to \infty$. $x(t) = o(y(t))$ if and only if, for every $\alpha > 0$, there exists a real-valued constant t_0 such that $|x(t)| \le \alpha |y(t)|$ for all $t \ge t_0$. $x(t) = \omega(y(t))$ if and only if, for every $\alpha > 0$, there exists a real-valued constant t_0 such that $|x(t)| \ge \alpha |y(t)|$ for all $t \ge t_0$. Graph $\mathcal{G}$ is an ordered pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consisting of a set of vertices (nodes) $\mathcal{V}$ and a set of edges $\mathcal{E}$. $(x, y, e_{x,y}) \in \mathcal{E}$ is an edge between nodes x and y with weight $e_{x,y}$. If the graph is directed, the edge is oriented from node x to node y; otherwise, the edge is not oriented. $\mathcal{G}(\mathbf{X})$ refers to an unweighted graph with unweighted adjacency matrix $\mathbf{X}$, whose non-zero elements are $[\mathbf{X}]_{ij} = 1$, and $\mathcal{G}(\bar{\mathbf{X}})$ refers to a weighted graph with weighted adjacency matrix $\bar{\mathbf{X}}$, whose non-zero elements are $[\bar{\mathbf{X}}]_{ij} = e_{i,j}$.

1.2.2 Theoretical Framework

We now develop the theoretical framework that motivates and guides this chapter. Consider an economic system with N total networked agents. Each agent i has a binary-valued attribute, b_i, with either $b_i = 0$ or $b_i = 1$; $n \le N$ agents have $b_i = 1$, the attribute's unit value. The global relative frequency of the attribute's unit value in the total population of agents is $f = \frac{n}{N}$. This quantity, f, is the economy's aggregate feature. In the rest of this chapter, we refer to f as the attribute's global relative frequency. Given f, there is a particular configuration, or arrangement, of the binary-valued attribute among agents in the economy. We define such a *configuration* as follows:

Definition 1.1 *A* configuration $\mathbf{b} \equiv \mathbf{b}(N, n)$ *of a binary-valued attribute in a population of N agents is an allocation of the attribute so that $b_i \in \{0, 1\}$ for all $i \in \{1, \dots, N\}$ and $\mathbf{b}^T \mathbf{1} = n$.*

A configuration $\mathbf{b} \equiv \mathbf{b}(N, n)$ of the binary-valued attribute among agents in the population is an allocation such that every agent has the attribute's zero or unit value, and the global relative frequency of the attribute in the population is $f = \frac{n}{N}$. We construct the $N \times 1$ configuration vector by taking each agent's attribute, b_i, and stacking this value for all agents in the population. From this vector, we can identify those agent indices with $b_i = 1$. Two configurations $\mathbf{b}$, $\mathbf{b}'$ are distinct if and only if $\mathbf{b} \neq \mathbf{b}'$ because the agent indices with $b_i = 1$ differ across these two configurations. We denote $\mathcal{B}(N, n)$ as the set of all possible configurations consistent with $f = \frac{n}{N}$, and the cardinality of this set is $|\mathcal{B}(N, n)| = \binom{N}{n}$.

Agents in this setting interact, and a network and its accompanying adjacency matrices capture these patterns of interaction.[2] The $N \times N$ unweighted adjacency matrix $\mathbf{A}$ captures the existence of linkages among agents; $[\mathbf{A}]_{ij} = 1$ if there is an edge from agent i to agent j. Meanwhile, the $N \times N$ weighted adjacency matrix $\bar{\mathbf{A}}$ captures the weights that agents assign to these linkages. $[\bar{\mathbf{A}}]_{ij} = e_{i,j}$ if the network has edge $(i, j, e_{i,j})$ from agent i to agent j with edge weight $e_{i,j}$. Agents allocate non-negative weight to each of their linkages, with the total weight allocated by a particular agent summing to 1; for an agent i, we therefore have $\sum_{j=1}^{N} [\bar{\mathbf{A}}]_{ij} = 1$.[3] This summation holds for all agents $i \in \{1, \ldots, N\}$, which makes matrix $\bar{\mathbf{A}}$ row-stochastic.

We are interested in the local relative frequency of the attribute's unit value, $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, for every configuration $\mathbf{b}(N, n) \in \mathcal{B}(N, n)$. In the rest of this work, we refer to $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ as the attribute's local relative frequency. This quantity depends on agents' interaction structure, $\bar{\mathbf{A}}$, it depends on the global frequency, n, of the attribute's unit value in the population of size N, and it depends on which

[2] In this chapter, networks can feature self-loops but not multiple edges.

[3] Chapters 2 and 3 of this dissertation relax the assumption that edge weights are non-negative and that the total allocated weight must sum to 1.

subset of agents on the network actually has that unit value, $\mathbf{b}$. Holding f fixed, as the configuration of the attribute adjusts and a different subset of agents has the attribute's unit value, we can imagine that $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ changes as well. Figure 1.1 plots the global and local relative frequencies of a binary-valued attribute across three different configurations. For all of these configurations, the attribute's global relative frequency is the same: $f = 0.50$. However, the local relative frequency of the attribute varies; it depends on which subset of agents actually has the attribute's unit value. For each configuration, this local relative frequency can also meaningfully deviate from the attribute's global relative frequency. For example, when agents 3 and 4 have the attribute, the local relative frequency of the attribute deviates positively, while when agents 1 and 2 have the attribute, the local relative frequency of the attribute deviates negatively. Scalar quantity $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is an abstract object; we are only able to directly compute this object once we assign to it a specific interpretation.

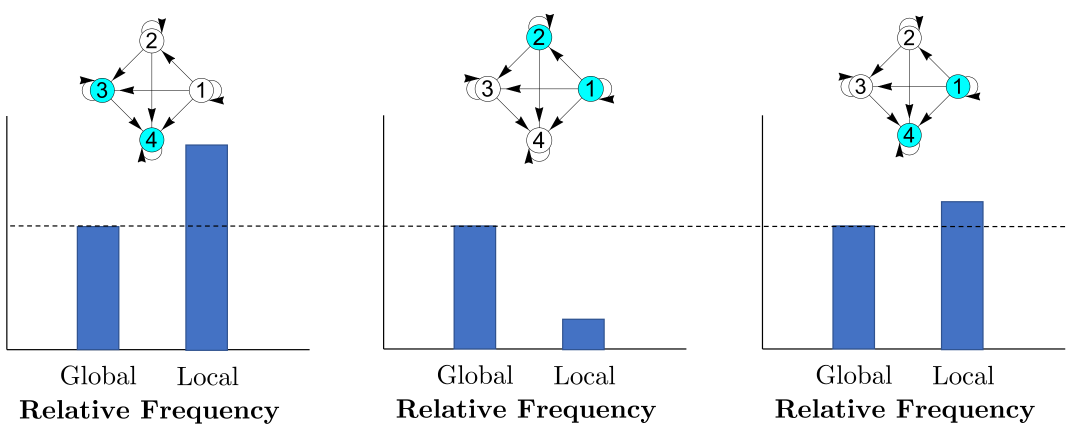


Figure 1.1: *The global relative frequency of the attribute and the local relative frequency of the attribute across three different configurations.*

Given $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ for every configuration $\mathbf{b}\left(N, n\right) \in \mathcal{B}\left(N, n\right)$, we then construct the distribution of possible local relative frequencies of the attribute. We define random variable $X(\bar{\mathbf{A}}, N, n)$ whose realizations are the configuration-specific quantities $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$. We are interested in the distributional features of $X(\bar{\mathbf{A}}, N, n)$. $G_{X(\bar{\mathbf{A}},N,n)}(t)$ is the CDF of $X(\bar{\mathbf{A}}, N, n)$ and $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is the PMF of $X(\bar{\mathbf{A}}, N, n)$. If every configuration is equally likely to occur:

$$G_{X(\bar{\mathbf{A}},N,n)}(t) = \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n) \in \mathcal{B}(N,n)} \mathbb{1}_{x(\bar{\mathbf{A}},\mathbf{b},N,n) \leq t}, \tag{1.1}$$

where $G_{X(\bar{\mathbf{A}},N,n)}(t)$ represents the fraction of configurations for which $x(\bar{\mathbf{A}}, \mathbf{b}, N, n) \leq t$. $G_{X(\bar{\mathbf{A}},N,n)}(t)$ is the precursor distribution from which we then proceed to construct the distribution of possible outcomes for the economy.

If the total number of configurations, $|\mathcal{B}(N,n)|$, is small, we can construct $G_{X(\bar{\mathbf{A}},N,n)}(t)$ given f by computing $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ configuration by configuration. In general, though, for $|\mathcal{B}(N,n)|$ small or large, we can construct $G_{X(\bar{\mathbf{A}},N,n)}(t)$ by decomposing $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ into two constituent quantities:

$$x(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N, n).$$

The first quantity is a fixed, network-derived vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$, and the second quantity is the particular configuration $\mathbf{b}(N,n)$ of the attribute among agents. The topology of the underlying network determines the values of $\mathbf{w}(\bar{\mathbf{A}})$. We can think of $\mathbf{w}(\bar{\mathbf{A}})$ as a vector that captures each agent's effective representation in the population. The higher an agent's weight, the higher the attribute's local relative frequency in the population when that agent possesses the attribute's unit value. To further see how $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ decomposes into $\mathbf{w}(\bar{\mathbf{A}})$ and $\mathbf{b}(N,n)$, note the following brief example: Suppose that we are interested in the relative frequency of the attribute, $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, for agent j in his immediate network neighborhood.

We have that

$$x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \frac{1}{|\mathcal{N}^{+}(j)|} \sum_{i \in \mathcal{N}^{+}(j)} b_i, \tag{1.2}$$

where $\mathcal{N}^{+}(j)$ is agent j's out-neighborhood on the network, that is, $\mathcal{N}^{+}(j)$ is the set of agents $i \in \{1, \ldots, N\}$ for which $[\bar{\mathbf{A}}]_{ji} > 0$. Re-writing this expression in Equation 1.2, we have:

$$x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \sum_{i=1}^{N} \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i b_i, \tag{1.3}$$

with $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i = 0$ if $i \notin \mathcal{N}^{+}(j)$ and $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i = \frac{1}{|\mathcal{N}^{+}(j)|}$ if $i \in \mathcal{N}^{+}(j)$. Agents not in agent j's neighborhood receive zero weight, while agents in agent j's neighborhood receive equal positive weight; the total weight allocated across all agents sums to 1.

The decomposition of $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ in Equation 1.3 indeed holds more generally, where we assume that each individual agent's weight is non-negative and $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{1} = 1$. Depending on the particular setting and the particular interpretation of $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ gets derived differently. However, it is this fixed network-derived vector of agent weights coupled with the combinatorially many possible configurations $\mathbf{b}\left(N, n\right) \in \mathcal{B}\left(N, n\right)$, given f, from which we can construct and compute the precursor distribution, $G_{X(\bar{\mathbf{A}},N,n)}\left(t\right)$. This decomposition allows us to compute the distributional features of $X\left(\bar{\mathbf{A}}, N, n\right)$ and the CDF $G_{X(\bar{\mathbf{A}},N,n)}\left(t\right)$ even when $|\mathcal{B}\left(N, n\right)| = \binom{N}{n}$ is large.

Figure 1.2 illustrates the theoretical framework that forms the basis for this chapter. We start off with an economic system that has a population of networked agents and an aggregate feature, namely the global relative frequency of the attribute. We assign a weight to each agent in the population and therefore derive from the underlying network a vector of agent weights. There are combinatorially many possible configurations of the attribute consistent with the system's aggregate feature.

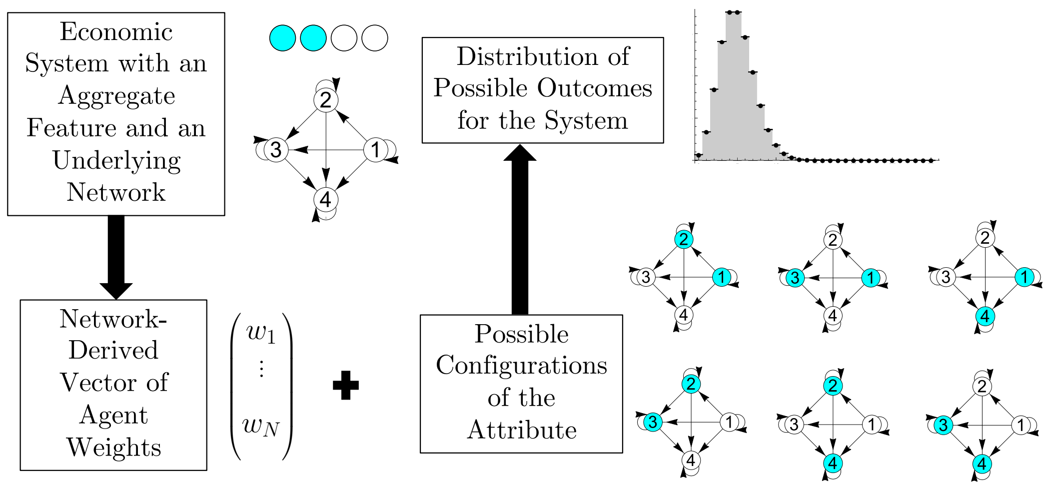


Figure 1.2: *Theoretical framework.*

We can compute the local relative frequency of the attribute for each configuration, given agents' weights. We then construct our precursor distribution of possible local relative frequencies of the attribute, and given agents' decision-making behavior, we construct the distribution of possible outcomes for the economic system.

1.2.3 Two Examples

We now walk through two examples that make both the theoretical framework and the quantities of interest more precise. Let's begin by assigning an interpretation to the binary-valued attribute. Let the binary-valued attribute denote employment status, with agent i unemployed if $b_i = 1$, and $b_i = 0$ otherwise. The global unemployment rate is $f = \frac{n}{N}$; in the language from before, f is the global relative frequency of the unemployment attribute. The agents are organized on a social observation network $G(\bar{\mathbf{A}})$ from which they observe each other's employment statuses and assign a weight to each of their observations. We are interested in determining the distribution of possible average local unemployment rates. This distribution of average local unemployment rates is a real-world manifestation of the distribution of possible local relative frequencies of the attribute from the previous subsection. We are examining possible local relative frequencies of the unemployment attribute, and more specifically, possible population-averaged local relative frequencies of the unemployment attribute for a given global unemployment rate.

Given the particular configuration of unemployment in the economy, we compute the average local unemployment rate, $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, as follows:

$$\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \frac{1}{N}\mathbf{1}^T\widehat{\mathbf{f}}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \frac{1}{N}\mathbf{1}^T\bar{\mathbf{A}}\mathbf{b}\left(N, n\right) = \left[\mathbf{d}_{\overline{w}}^{-}\left(\bar{\mathbf{A}}\right)\right]^T\mathbf{b}\left(N, n\right).$$

$\widehat{\mathbf{f}}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is the $N \times 1$ population vector of agents' local unemployment rates. Each agent's local unemployment rate is calculated by determining the weighted relative frequency of the unemployment attribute in that agent's immediate out-neighborhood. The local unemployment rate for agent i is therefore $\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\bar{\mathbf{A}}\right]_{i*}\mathbf{b}\left(N, n\right)$, which makes the population vector of local unemployment rates $\widehat{\mathbf{f}}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \bar{\mathbf{A}}\mathbf{b}\left(N, n\right)$. The relevant network-derived vector of agent weights is $\mathbf{d}_{\overline{w}}^{-}\left(\bar{\mathbf{A}}\right) = \frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1}$, the vector of average weighted in-degrees. Note the parallel between the decomposition of $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, the average local unemployment rate, and $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, the local relative frequency of the attribute:

$$\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{d}_{\overline{w}}^{-}\left(\bar{\mathbf{A}}\right)\right]^T\mathbf{b}\left(N, n\right) \text{ and}$$
$$x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T\mathbf{b}\left(N, n\right).$$

Here, the random variable of interest is $\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right)$ with configuration-specific realization $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, and CDF $G_{\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right)}\left(t\right)$ and PMF $g_{\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right)}\left(t\right)$, while in the previous subsection, the random variable of interest was $X\left(\bar{\mathbf{A}}, N, n\right)$ with configuration-specific realization $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, and CDF $G_{X\left(\bar{\mathbf{A}}, N, n\right)}\left(t\right)$ and PMF $g_{X\left(\bar{\mathbf{A}}, N, n\right)}\left(t\right)$. There is an exact parallel between the computation of $G_{\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right)}\left(t\right)$

and $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, and the respective computation of $G_{X(\bar{\mathbf{A}},N,n)}(t)$ and $g_{X(\bar{\mathbf{A}},N,n)}(t)$. When every configuration is equally likely to occur,

$$G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t) = \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\widehat{f}_{avg}(\bar{\mathbf{A}},\mathbf{b},N,n)\leq t}$$

which exactly parallels the expression for $G_{X(\bar{\mathbf{A}},N,n)}(t)$ in Equation 1.1.

We can now proceed to our first example, in which we study the relationship between agents' interaction network and the distribution of possible average local unemployment rates:

Example 1.1 (Average Local Unemployment Rate, N = 4) *Consider an economy with $N = 4$ agents and an unemployment rate of $f = 0.25$. Agents' social observation network, from which they observe each others' employment statuses, is depicted in Figure 1.3. The corresponding row-stochastic weighted adjacency matrix, $\bar{\mathbf{A}}$, is immediately below. Assuming that each configuration of unemployment in the economy is equally likely, we can compute the distribution, $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, of possible average local unemployment rates:*

$$\bar{\mathbf{A}} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t) = \begin{cases} 0.25 & \text{if } \widehat{F}_{avg}(\bar{\mathbf{A}},N,n) = 0.0625 \\ 0.25 & \text{if } \widehat{F}_{avg}(\bar{\mathbf{A}},N,n) \approx 0.146 \\ 0.25 & \text{if } \widehat{F}_{avg}(\bar{\mathbf{A}},N,n) \approx 0.271 \\ 0.25 & \text{if } \widehat{F}_{avg}(\bar{\mathbf{A}},N,n) \approx 0.521 \end{cases}.$$

In this example, there are four configurations of unemployment in the economy consistent with $f = 0.25$: $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and $\mathbf{e}_4 \in \mathbb{R}^4$. For each configuration, given that $\widehat{\mathbf{f}}(\bar{\mathbf{A}},\mathbf{b},N,n) = \bar{\mathbf{A}}\mathbf{b}(N,n)$, we can compute the average local unemployment rate:

For $\mathbf{b} = \mathbf{e}_1$, $\widehat{\mathbf{f}}(\bar{\mathbf{A}},\mathbf{b},N,n) = \begin{pmatrix} 0.25 & 0 & 0 & 0 \end{pmatrix}^T$ and $\widehat{f}_{avg}(\bar{\mathbf{A}},\mathbf{b},N,n) = 0.0625$.

For $\mathbf{b} = \mathbf{e}_2$, $\widehat{\mathbf{f}}(\bar{\mathbf{A}},\mathbf{b},N,n) = \begin{pmatrix} 0.25 & 0.33 & 0 & 0 \end{pmatrix}^T$ and $\widehat{f}_{avg}(\bar{\mathbf{A}},\mathbf{b},N,n) \approx 0.146$.

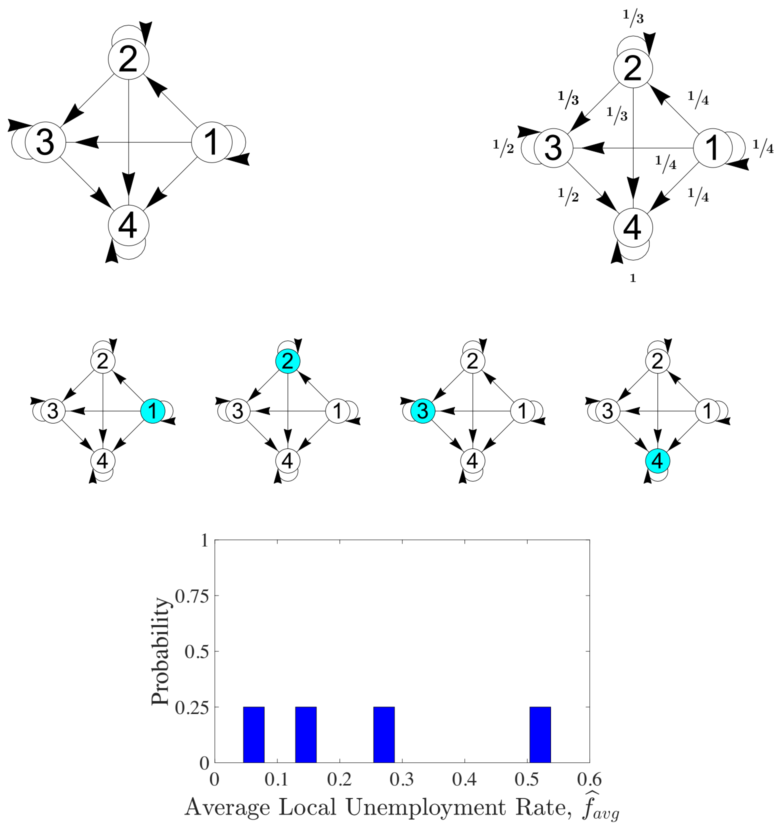


Figure 1.3: *Graphs $\mathcal{G}(\mathbf{A})$ (top left) and $\mathcal{G}(\bar{\mathbf{A}})$ (top right) from Example 1.1. The four possible configurations, $\mathbf{b}(N,n)$, of the binary-valued attribute given that $f = 0.25$ (middle). The probability distribution, $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, of possible average local unemployment rates given that $f = 0.25$ (bottom).*

For $\mathbf{b} = \mathbf{e}_3$, $\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \begin{pmatrix} 0.25 & 0.33 & 0.50 & 0 \end{pmatrix}^T$ and $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \approx 0.271$.

For $\mathbf{b} = \mathbf{e}_4$, $\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \begin{pmatrix} 0.25 & 0.33 & 0.50 & 1.00 \end{pmatrix}^T$ and $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \approx 0.521$.

From these values, we can then construct the probability distribution $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, which is depicted in Figure 1.3. This probability distribution is our main object of interest, and most of the present work is devoted to understanding and characterizing this type of object.

The topology of agents' observation network in Example 1.1 generates wide variation in individual agents' local unemployment rates for a particular configuration of unemployment in the economy, and it generates variation in individual agents' local unemployment rates across configurations. The latter variation arises from agents having different effective representations in the population, or in this particular setting, different levels of observability. Agent 1 has a low average weighted in-degree, and therefore a low agent weight and poor observability, so when he is unemployed, agents 1, 2, 3, and 4 respectively have local unemployment rates of 25 percent, 0 percent, 0 percent, and 0 percent; agent 4 meanwhile has a high average weighted in-degree, and therefore a high agent weight and strong observability, so when he is unemployed, agents 1, 2, 3, and 4 respectively have local unemployment rates of 25 percent, 33 percent, 50 percent, and 100 percent. This variation in the observability of agents causes the average local unemployment rate to change with configuration. When agent 1 is unemployed, the average local unemployment rate is 6.25 percent, while when agent 4 is unemployed, the average local unemployment rate is 52.1 percent. These values for the average local unemployment rate also strongly deviate from the actual unemployment rate of 25 percent. If agents form macroeconomic sentiment from their local rates of unemployment, then depending on the particular configuration of unemployment in the economy, agents on average might feel that the economy is doing much better or worse than its fundamentals would otherwise suggest. If the outcome of the economy somehow depends on this average local unemployment rate, and this average local unemployment rate substantially varies with the particular configuration of unemployment in the economy, then we consider the economy to be strongly configuration-dependent.

We can trace different pathways for the average local unemployment rate as the global unemployment rate evolves. Figure 1.4 features four potential pathways for the average local unemployment rate; Figure 1.4 also plots, in the background,

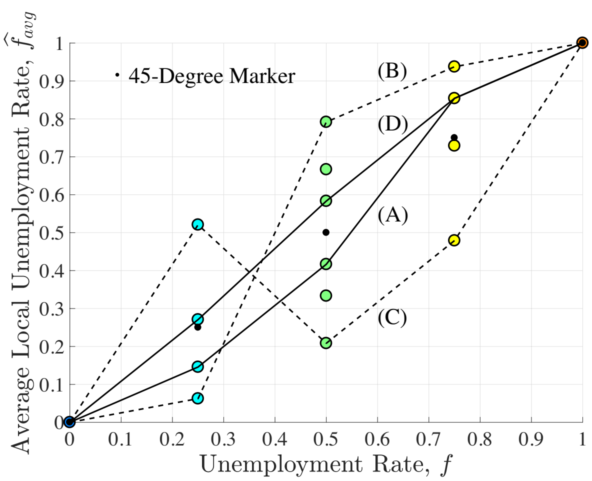


Figure 1.4: *Possible average local unemployment rates,* $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$*, given the global unemployment rate,* f*, and potential pathways (A-D) for the average local unemployment rate as the global unemployment rate evolves.*

the set of all possible average local unemployment rates for each feasible level of unemployment in the economy. Changes in configuration can accommodate different phenomena that would otherwise not emerge from just the aggregate properties of the system. Path (B) illustrates how there can be dramatic swings in sentiment for a small adjustment to the the global unemployment rate, the system's aggregate feature; as the unemployment rate increases from 25 percent to 50 percent, there is a 72.9-percentage-point increase in the average local unemployment rate. Path (C) illustrates how sentiment can move in a direction opposite to that of fundamentals. Even though the unemployment rate is declining from 50 percent to 25 percent, the average local unemployment rate increases 31.3 percentage points. The decrease in unemployment would suggest that the economy is improving, but the agents in the population on average locally observe the economy to be worsening. Paths (A) and (D) illustrate hysteresis within the economy. Suppose that the economy takes path (A) as the unemployment rate increases and it takes path (D) as the unemployment rate decreases. Even though the economy is traversing the same set of unemployment rates, it can experience different average local unemployment rates. Configuration dependence of the economic system enables the existence of such phenomena.

In the second example, we study the distribution of possible average local unemployment rates, $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, in a setting with a larger sample population:

Example 1.2 (Average Local Unemployment Rate, N = 15) *Consider an economy with* $N = 15$ *agents and an unemployment rate of* $f = 0.20$*. Agents' social observation network,* $\mathcal{G}(\mathbf{A})$*, is formed from preferential attachment, with a self-loop for every node (see Figure 1.5). Assuming that agents equally weight each of their observations and each configuration of unemployment in the economy is equally likely, the distribution of possible average local unemployment rates,* $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$*, is depicted at the bottom right of Figure 1.5.*

In this second example, there are $\binom{15}{3} = 455$ possible configurations of the unemployment attribute among agents in the population consistent with a 20-percent unemployment rate. As in the first example, we compute the average local unemployment rate configuration by configuration, and we then construct the accompanying probability distribution. We observe substantial heterogeneity in agents' average weighted in-degrees, $\mathbf{d}_w^-(\bar{\mathbf{A}}) = \frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1}$, as can be viewed in the bottom left of Figure 1.5. With such heterogeneity in agents' weights, individual agents' contributions to the average local unemployment rate measurably differ when they become unemployed. As a result, the probability distribution of possible average local unemployment rates has sizable variance. For a 20-percent unemployment rate, the average local unemployment rate can vary from 11.9 percent to 33.1 percent. If the local unemployment rate is relevant for agent decision-making, the economy is strongly dependent on the underlying configuration of unemployment in the population.

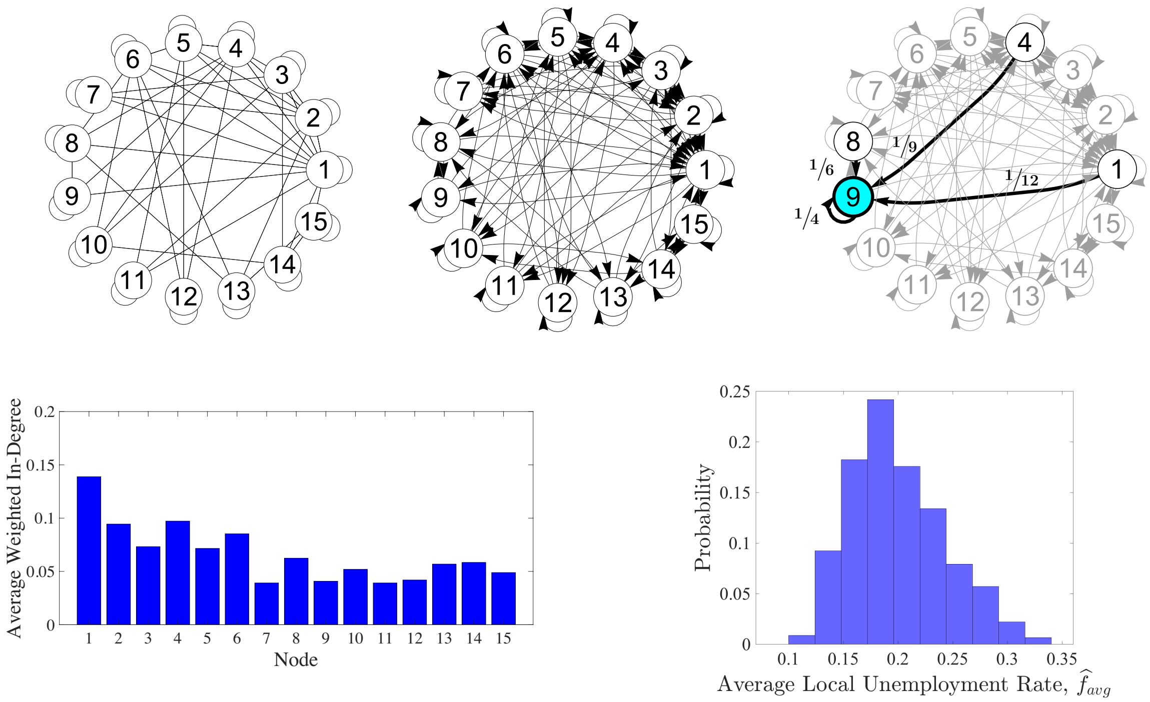


Figure 1.5: *Graphs* $\mathcal{G}(\mathbf{A})$ *(top left) and* $\mathcal{G}(\bar{\mathbf{A}})$ *(top center) for Example 1.2. Calculating a node's weighted in-degree (top right) and the plot of average weighted in-degrees,* $\mathbf{d}_w^-(\bar{\mathbf{A}})$ *(bottom left). The distribution of average local unemployment rates,* $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$*, for* $f = 0.20$ *(bottom right).*

We can construct such a probability distribution of possible average local unemployment rates for every feasible level of unemployment in the economy. We can conceivably compute the average local unemployment rate configuration by configuration for a given global unemployment rate, and we can then plot the corresponding probability distribution. The top of Figure 1.6 plots the set of possible average local unemployment rates for a fixed global unemployment rate, and the bottom of Figure 1.6 plots the corresponding probability distribution of possible average local unemployment rates for each feasible global level of unemployment in the economy. When one agent is unemployed, there are just 15 possible configurations, and when 7 or 8 agents are unemployed, there are 6435 possible configurations. We continue to observe strong configuration dependence of

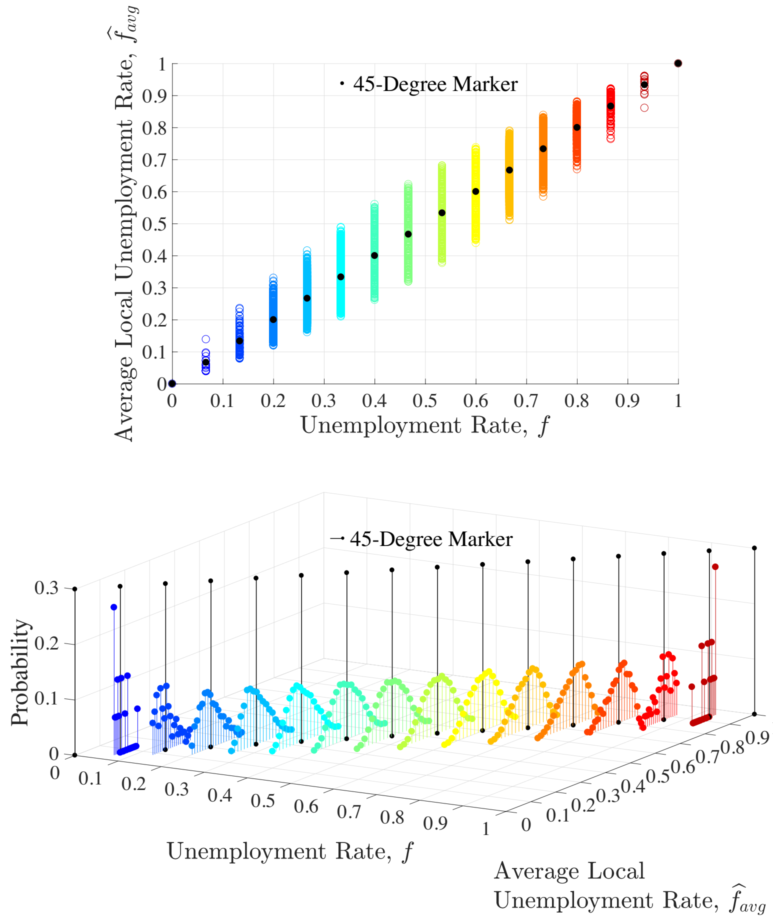


Figure 1.6: *Possible average local unemployment rates, $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, for a given global unemployment rate, f, (top) and the probability distribution of possible average local unemployment rates, $g_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, for each global level of unemployment in the economy (bottom).*

the system, and the average local unemployment rate can substantially deviate from the actual global unemployment rate depending on the particular configuration of unemployment in the economy. As the population size gets larger, the total number of possible configurations, $\binom{N}{n}$, grows combinatorially, and it becomes less feasible to construct the distribution of possible average local unemployment rates configuration by configuration. In Section 1.6, we present a set of theoretical results that allows us to construct this probability distribution and compute its features in closed form for every population size, no matter how large the population size gets.

We can imagine that an agent's local unemployment rate is a determinant of his or her sentiment about the macroeconomy, and the average local unemployment rate is a population-wide indicator of macroeconomic sentiment. Such macroeconomic sentiment can influence agents' behavior. In the next section, we consider a setting in which the local unemployment rate impacts agents' voting decision. Depending on the topology of agents' observation network, the overall election outcome, and thus the outcome for the economy, can be strongly configuration dependent.

1.3 Macroeconomic Sentiment and Election Outcomes

We use the theoretical framework that we just developed to examine election outcomes in a stylized setting. We study a population of voters who must choose between two candidates named Hillary Clinton and Donald Trump. Each voter's macroeconomic sentiment, formed from that voter's local unemployment rate, influences his or her voting decision, so configuration-induced variations in sentiment for fixed macroeconomic fundamentals can alter individual agents' voting decisions, and thereby alter the election outcome and the outcome of the economy. This section essentially shows how we can embed our theoretical framework developed in the previous section into a real-world setting and extract economic meaning.

1.3.1 Model

We begin by considering a population of N agents, each of whom is a voter in the election. Each agent faces a binary choice problem: either vote for Hillary Clinton or Donald Trump. To make their decisions, agents consider the candidates' policies. There are P issues, and the candidates construct a policy for every issue. For each policy put forth by a particular candidate, there is an associated scalar benefit. Scalar benefits are separable across issues. The $P \times 1$ vectors of benefits corresponding to the policies of Clinton and Trump are respectively $\mathbf{x}_C$ and $\mathbf{x}_T$. Agent i weights each candidate's set of policies using the $P \times 1$ weighting vector $\boldsymbol{\alpha}_i$, with $\boldsymbol{\alpha}_i^T \mathbf{1} = 1$. $\boldsymbol{\alpha}_i \neq \boldsymbol{\alpha}_j$ for agents i and j represents preference heterogeneity. Agent i votes for Clinton if $u_i(\mathbf{x}_C, \boldsymbol{\alpha}_i) > u_i(\mathbf{x}_T, \boldsymbol{\alpha}_i)$, and agent i votes for Trump if $u_i(\mathbf{x}_T, \boldsymbol{\alpha}_i) > u_i(\mathbf{x}_C, \boldsymbol{\alpha}_i)$.

The utilities that agent i respectively receives from the implementation of Clinton's policies and Trump's policies are:

$$u_i(\mathbf{x}_C, \boldsymbol{\alpha}_i\,; \epsilon_{iC}) = \boldsymbol{\alpha}_i^T \mathbf{x}_C + \epsilon_{iC} \quad \text{and} \quad u_i(\mathbf{x}_T, \boldsymbol{\alpha}_i\,; \epsilon_{iT}) = \boldsymbol{\alpha}_i^T \mathbf{x}_T + \epsilon_{iT},$$

where ϵ_{iC} and ϵ_{iT} are agent-specific, choice-specific shocks, as in Heckman and Snyder (1997), and $E[u_i(\mathbf{x}_\ell, \boldsymbol{\alpha}_i\,; \epsilon_{i\ell})] = \boldsymbol{\alpha}_i^T \mathbf{x}_\ell$ for $\ell \in \{C, T\}$. Subutilities $\boldsymbol{\alpha}_i^T \mathbf{x}_C$ and $\boldsymbol{\alpha}_i^T \mathbf{x}_T$ both have a linear form; each one represents the weighted sum of benefits accrued from the policies of a particular candidate.

Specifying the decision-making rule for each agent, let π_{iT} be the probability that agent i votes for Trump given $\mathbf{x}_T$, $\mathbf{x}_C$, and $\boldsymbol{\alpha}_i$, and let $\pi_{iC} = 1 - \pi_{iT}$ be the probability that agent i votes for Clinton. Shocks ϵ_{iT}, ϵ_{iC} are assumed to be both independent of the candidates' policies and independent across voters. Define $\eta_i = \epsilon_{iT} - \epsilon_{iC}$, which is independent of $\boldsymbol{\alpha}_i^T \mathbf{x}_T - \boldsymbol{\alpha}_i^T \mathbf{x}_C$. Setting $\eta_i \overset{\text{iid}}{\sim} \text{Uniform}(-\beta, \beta)$ for every agent $i \in \{1, \ldots, N\}$,

$$\pi_{iT} = \frac{1}{2} + \frac{\boldsymbol{\alpha}_i^T (\mathbf{x}_T - \mathbf{x}_C)}{2\beta} \quad \text{and} \quad \pi_{iC} = \frac{1}{2} + \frac{\boldsymbol{\alpha}_i^T (\mathbf{x}_C - \mathbf{x}_T)}{2\beta},$$

where $\pi_{iT} = \Pr[u_i(\mathbf{x}_T, \boldsymbol{\alpha}_i\,; \epsilon_{iT}) > u_i(\mathbf{x}_C, \boldsymbol{\alpha}_i\,; \epsilon_{iC})]$. If both candidates' policies confer the same benefits, $\mathbf{x}_T = \mathbf{x}_C$ and agent i is equally likely to vote for either candidate: $\pi_{iT} = \pi_{iC} = 0.5$.[4]

[4]We restrict the allowable values for β. We set β so that $\pi_{iT}, \pi_{iC} \in (0, 1)$ for all $i \in \{1, \ldots, N\}$.

Let the first policy concern jobs and unemployment. We assume that the weight an agent assigns to this policy directly depends on his or her local unemployment rate, $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$:

$$[\boldsymbol{\alpha}_i]_1 = \alpha_{i1} = \widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\,;$$

$\mathbf{b} \equiv \mathbf{b}(N, n)$ is the configuration of unemployment in the economy, with $[\mathbf{b}]_i = 1$ if agent i is unemployed and $[\mathbf{b}]_i = 0$ otherwise, and the overall unemployment rate is $f = \frac{n}{N}$. The higher an agent's local unemployment rate, the more that agent cares about the issue of jobs and unemployment in the economy. This is consistent with the findings of Bisgaard, Dinesen, and Sønderskov (2016), who observe that Danish voters' dissatisfaction with the national economy increases with the local unemployment rate, defined as the fraction of all unemployed residents within a fixed meter radius from the voter's place of residence. The work of Healy and Lenz (2017) meanwhile justifies the dependence of agents' voting decisions on the local unemployment rate, as these authors demonstrate that local unemployment conditions impact the national voting outcome.

Therefore, if $\boldsymbol{\alpha}_i^T \mathbf{x}_T > \boldsymbol{\alpha}_i^T \mathbf{x}_C$, we want it to be possible for $u_i(\mathbf{x}_C, \boldsymbol{\alpha}_i; \epsilon_{iC}) > u_i(\mathbf{x}_T, \boldsymbol{\alpha}_i; \epsilon_{iT})$, and if $\boldsymbol{\alpha}_i^T \mathbf{x}_C > \boldsymbol{\alpha}_i^T \mathbf{x}_T$, we want it to be possible for $u_i(\mathbf{x}_T, \boldsymbol{\alpha}_i; \epsilon_{iT}) > u_i(\mathbf{x}_C, \boldsymbol{\alpha}_i; \epsilon_{iC})$. For this to happen, we need there to exist separate realizations of η_i so that

$$\eta_i < \boldsymbol{\alpha}_i^T(\mathbf{x}_C - \mathbf{x}_T) \text{ and } \eta_i > \boldsymbol{\alpha}_i^T(\mathbf{x}_C - \mathbf{x}_T).$$

Since $\eta_i \overset{\text{iid}}{\sim} \text{Uniform}(-\beta, \beta)$ for every agent $i \in \{1, \ldots, N\}$, it follows that we must set

$$\beta > \max_{i \in \{1, \ldots, N\}} \left| \boldsymbol{\alpha}_i^T(\mathbf{x}_T - \mathbf{x}_C) \right|.$$

With $\pi_{iT} \in (0, 1)$, we are able to aggregate individual agents' decision-making functions, π_{iT}, by simple summation of π_{iT} across agents.

The aggregate equation for the expected fraction of votes for Trump is now:

$$\frac{1}{N}\sum_{i=1}^{N}\pi_{iT} = \frac{1}{2} + \frac{1}{2\beta}\left[\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\left(x_{T,1} - x_{C,1}\right) + \frac{\sum_{i=1}^{N}\alpha_{i2}}{N}\left(x_{T,2} - x_{C,2}\right) + \cdots \right.$$
$$\left. + \frac{\sum_{i=1}^{N}\alpha_{iP}}{N}\left(x_{T,P} - x_{C,P}\right)\right].$$

The overall weight accorded to the issue of jobs and unemployment in the economy is the average local unemployment rate. If this quantity varies enough with the configuration of unemployment in the economy, then we can potentially anticipate different voting outcomes for a particular global unemployment rate.

Let's assume that Trump's jobs policy confers a greater benefit to voters than Clinton's jobs policy: $x_{T,1} > x_{C,1}$ with $x_{T,1} - x_{C,1} = 9$. Let's also assume that $x_{C,j} - x_{T,j} = x_{C,\ell} - x_{T,\ell} = 1$ for all $j, \ell \neq 1$, so the policy put forth by Clinton for every other issue yields a benefit that exceeds that of Trump's corresponding policy. A higher average local unemployment rate favors the election of Trump. The aggregate equation for the expected fraction of votes for Trump becomes:

$$\frac{1}{N}\sum_{i=1}^{N}\pi_{iT} = \frac{1}{2} + \frac{1}{2\beta}\left[10\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) - 1\right]. \tag{1.4}$$

If the average local unemployment rate is 10 percent, either candidate is equally likely to win the election. If the average local unemployment rate exceeds 10 percent, the expected vote share for candidate Trump exceeds 50 percent, so the voting outcome favors Trump. If the average local unemployment rate is less than 10 percent, the expected vote share for candidate Trump is less than 50 percent, so the voting outcome favors Clinton. We proceed to construct voters' observation network so that we can determine the possible values of the average local unemployment rate consistent with the economy's overall unemployment rate. Voters compute their local unemployment rates from their nodes on this network.

These local unemployment rates affect individual voting behavior, and the average local unemployment rate affects aggregate voting behavior and the election outcome.

1.3.2 Constructing Voters' Observation Network

To construct voters' observation network, we draw on details from the 2016 U.S. presidential election. The observation network has 137.5 million nodes, the total number of voters in the 2016 U.S. presidential election.[5] Each node has a self-loop because voters observe their own employment statuses. We assume that each agent has, on average, 50 reciprocal linkages; these linkages may be formed with relatives, colleagues, acquaintances, and so on. We accordingly construct an undirected Erdös-Rényi graph whose expected degree is 50. We refer to this network of 137.5 million nodes with its self-loops and Erdös-Rényi linkages as the base graph.

In addition to the linkages comprising the base graph, we introduce media-originating directed linkages. During the weeks and months preceding the election, we assume that voters engage with a variety of news/talk media outlets that feature stories about employed and unemployed individuals. These featured individuals can shift voters' perceptions of unemployment in the economy.[6] We therefore gather statistics on television network viewership, radio show listenership, and newspaper, magazine, business journal, and business publication circulation in the United States.

Appendix A.1 provides more details about viewership, listenership, and readership statistics for media in the United States and overall data construction.

[5] "Black voter turnout fell in 2016, even as a record number of Americans cast ballots," Pew Research Center, May 12, 2017.

[6] Indeed, Goidel and Langley (1995) and Hetherington (1996) study the role of the media in influencing voters' evaluations of the macroeconomy.

There is a total of 1867 different news/talk media sources. As observed in Figure 1.7, audience sizes are heavy-tailed. We assume that news/talk media outlets feature an employed or unemployed individual in one story per week for 15 weeks.[7] Therefore, to construct the network of media-originating linkages, for each media outlet, we randomly select the set of audience members and we randomly select the set of 15 individuals that are featured in that news outlet's stories. Directed edges are then drawn from the set of audience members to each featured agent. This method of network construction is carried out for all 1867 news/talk media sources.

Appendix A.1 presents detailed summary statistics for the base graph, the media-originating graph, and the composite graph that pools both base and media-originating linkages. Figure 1.7 plots the counter-cumulative distribution function of degrees for the base graph, the counter-cumulative distribution functions of out-degrees and in-degrees arising from the network of media-based linkages, and the counter-cumulative distribution functions of out-degrees and in-degrees arising from the composite network. For the base graph, with its 3,575,017,297 undirected edges, the average degree is 51.0 with a standard deviation of 7.07. We obtain this average degree because the Erdös-Rényi graph has an average degree of 50 and each agent has a self-loop. In the media-originating graph, with its 2,712,493,694 directed edges, the average out-degree is 19.7 with a standard deviation of 17.0, and the average in-degree is 19.7 with a standard deviation of 8,633.3.[8] The counter-cumulative distribution function of out-degrees for the media-originating graph is a step function because agents accumulate 15 out-edges for every media source

[7] We choose 15 weeks because that is the number of weeks that elapsed from the conclusion of the Republican and Democratic National Conventions until Election Day for the 2016 U.S. presidential election. The Republican National Convention, Democratic National Convention, and Election Day respectively took place from July 18-21, 2016, from July 25-28, 2016, and on November 8, 2016.

[8] The average out-degree and the average in-degree for any directed graph take the same value since the total number of out-edges equals the total number of in-edges.

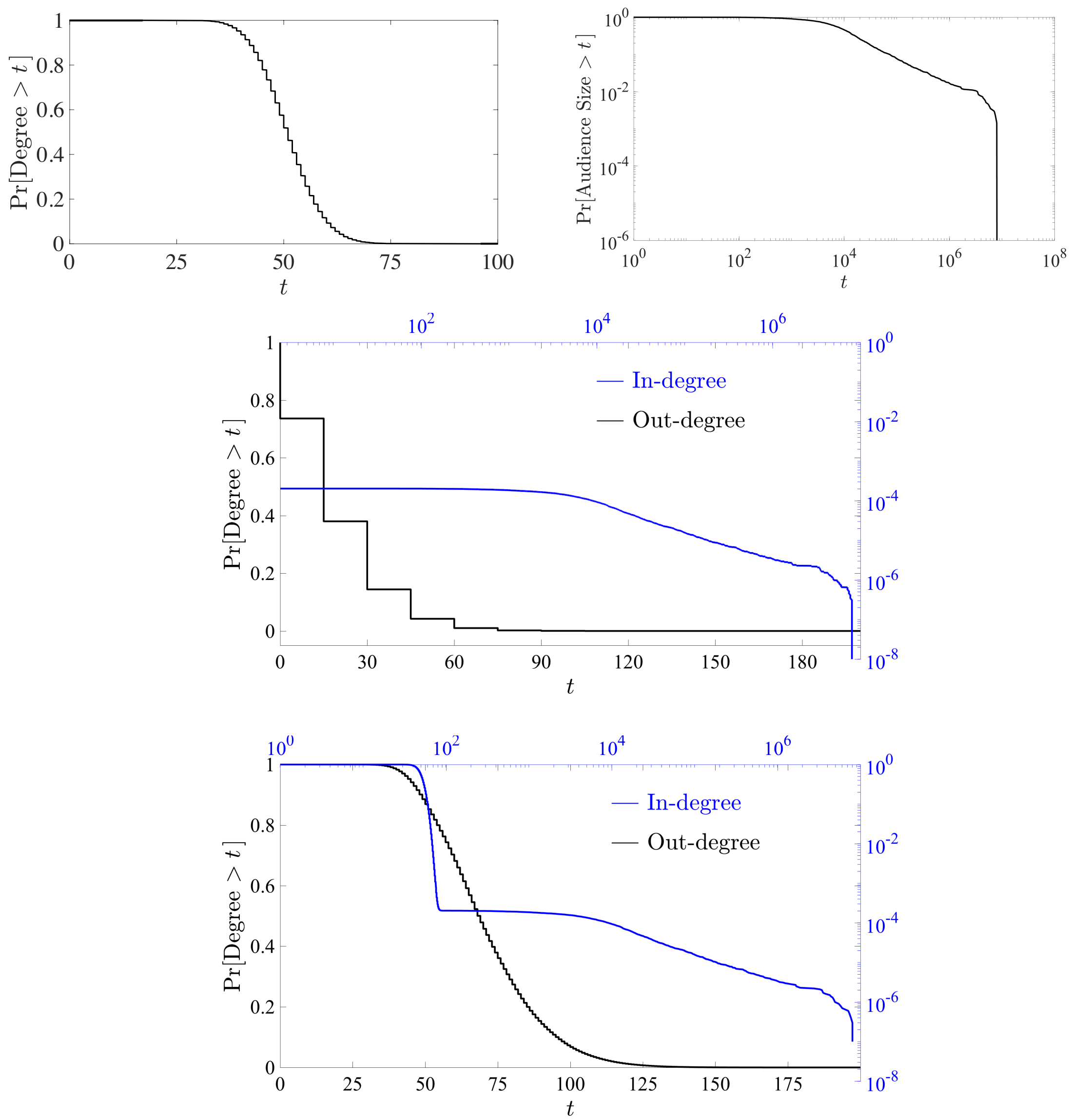


Figure 1.7: *Counter-cumulative distribution function (CCDF) of degrees for the base graph (top left). CCDF of audience sizes for 1867 different media sources (top right). CCDFs of out- and in-degrees for the network of media-originating linkages (middle). CCDFs of out- and in-degrees for the composite network (bottom).*

in which they are an audience member. Therefore, out-degrees for the media-originating graph occur in multiples of 15. Most voters have zero in-degree in the media graph because they are not featured in news/talk media outlets; there are only 28,003 individuals featured in employment-related news stories. The probability of a non-zero in-degree for this graph is 2.04×10^{-4}, which we can observe in Figure 1.7 (middle). Since the counter-cumulative distribution function of audience sizes across media sources is heavy-tailed, the counter-cumulative distribution function of in-degrees for the media-originating graph is similarly heavy-tailed.

In the composite graph, the average out-degree is 70.7 with a standard deviation of 18.4. As depicted in Figure 1.7, we see that the counter-cumulative distribution function of out-degrees for the composite graph takes the same shape as the counter-cumulative distribution function of out-degrees for the base graph, except that the former distribution is shifted to the right. Each agent has accumulated additional out-edges from media-originating linkages, which generates a shift in the distribution function. On average, the total number of media sources to which people are exposed is 1.32. For the composite graph, the average in-degree is 70.7, with a standard deviation of 8,633.3, and the maximum in-degree is 8,020,651. The counter-cumulative distribution function of in-degrees for the composite network (Figure 1.7, bottom) directly incorporates the distributional features of the counter-cumulative distribution function of degrees for the base graph (Figure 1.7, top left) and the counter-cumulative distribution function of in-degrees for the media graph (Figure 1.7, middle). Most agents are not featured by the media, so their in-degree is equal to their degree from the base graph. A small fraction of agents are featured in the media, so their in-degree is equal to their degree from the base graph plus their in-degree from the media graph. The counter-cumulative distribution function of in-degrees for the composite graph therefore becomes heavy-tailed.

1.3.3 When Configurations are Equally Likely: Distribution of Possible Average Local Unemployment Rates and the Expected Voting Outcome

We can now study the distribution of possible average local unemployment rates and the probability that the election outcome favors each individual candidate. We set the unemployment rate in this economy to 9.6 percent. We decide to use the October 2016 U-6 unemployment rate in the United States specified by the Bureau of Labor Statistics. This value is the national unemployment rate that immediately precedes the 2016 U.S. presidential election.[9] Since we are interested in capturing individuals' macroeconomic sentiment, we use the U-6 unemployment series because it counts those people who are discouraged workers or underemployed for economic reasons as unemployed.

We would like to determine the distribution of possible average local unemployment rates, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, in the economy given that there is an overall 9.6-percent unemployment rate. For each possible configuration of unemployment in the economy, the average local unemployment rate is computed as follows: $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{d}_w^-(\bar{\mathbf{A}})]^T \mathbf{b}(N, n)$. As in Section 1.2, the average local unemployment rate depends on the vector of average weighted in-degrees, $\mathbf{d}_w^-(\bar{\mathbf{A}}) = \frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1}$, of the underlying social observation network, $\mathcal{G}(\bar{\mathbf{A}})$, and the particular configuration of unemployment, $\mathbf{b}(N, n)$. The capacity for $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ to be meaningfully non-degenerate depends on the properties of $\mathbf{d}_w^-(\bar{\mathbf{A}})$, the latter of which is derived from the 137.5-million-node voter observation network.

Appendix A.2 considers the case in which agents' observation network solely

[9] The U-6 unemployment rate is defined as the "total unemployed, plus all marginally attached workers plus total employed part time for economic reasons, as a percent of all civilian labor force plus all marginally attached workers."

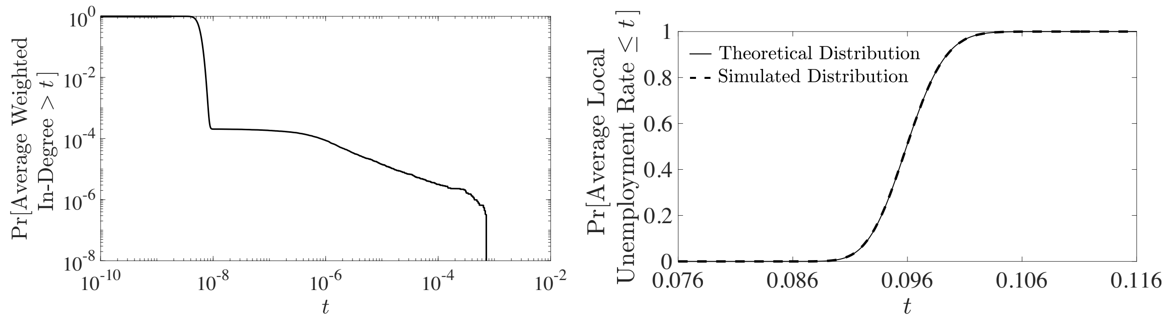


Figure 1.8: *Counter-cumulative distribution function of average weighted in-degrees for the composite network, assuming that each agent assigns an equal weight to each of his out-linkages (left). Distribution of the average local unemployment rate, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, when $f = 0.096$, assuming that configurations of unemployment in the economy are equally likely (right).*

consists of the base graph; in that setting, the average local unemployment rate does not meaningfully vary with configuration. Here, we take agents' observation network to be the more realistic composite graph. We compute each agent's average weighted in-degree by assuming that agents equally weight each of their observations of employment status. Once we compute this vector of agent weights, we can determine each agent's effective representation in the population. On average, each agent has an effective weight of 1 agent, which we would expect. The effective minimum weight is 0.249 agents, and the effective maximum weight is 98,733.1 agents. The median agent has an effective weight of 0.756 agents. The left side of Figure 1.8 plots the counter-cumulative distribution function of average weighted in-degrees. This distribution of agent weights is heavy-tailed. There is a relatively small subset of agents in the entire voting population that is particularly influential from being featured in the media.

We observe the distribution of possible average local unemployment rates, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, on the right side of Figure 1.8. For every distributional feature of $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ that we highlight in this section, there is a corresponding theorem presented in later sections of this work that shows how to compute that quantity in closed form. On the right side of Figure 1.8, the theoretical CDF for $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ overlays an empirical CDF.[10] The empirical CDF is constructed by randomly drawing 100,000 configurations of unemployment from the set of all possible configurations consistent with a 9.6-percent unemployment rate, and then computing the associated average local unemployment rate for each configuration. The theoretical and empirical mean of this distribution is 0.096.[11] The theoretical standard deviation for this distribution is 0.00266, or 0.266 percentage points, and the size of two standard deviations about the distribution's mean value is 1.07 percentage points.[12] Staying within this two-standard-deviation band, the average local unemployment rate can generally vary from 9.07 percent to 10.1 percent. The lowest possible average local unemployment rate is 5.53 percent, and the highest possible average local unemployment rate is 33.3 percent.[13] The average local unemployment rate here exhibits substantial configuration dependence, especially when we consider the magnitude of fluctuations in the actual unemployment rate over the course of a business cycle. For the most recent complete business cycle dated by the NBER, that is, from March 2001 until December 2007 (i.e., from peak to peak), the U-6 unemployment series varied from 7.3 percent to 10.4 percent, a difference of just 3.1 percentage points. Variations in the average local unemployment rate are indeed large enough that they can potentially mimic variations in business cycle conditions.

There are two different ways that we can understand the existence of such strong configuration dependence for this very-large-N economic system. First, strong variation in $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ emerges for a given f because the set of in-

[10]The theoretical CDF is constructed from Theorem 1.13 in Section 1.6.

[11]The theoretical mean is computed from Theorem 1.8 in Section 1.6.

[12]The theoretical standard deviation is computed from Theorem 1.9 in Section 1.6.

[13]Both quantities are constructed from Theorem 1.10 in Section 1.6.

degrees for the composite graph (Figure 1.7, bottom) has a very high variance. Second, strong variation in $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ emerges because the distribution of agent weights, that is, the distribution of average weighted in-degrees for the composite graph (Figure 1.8, left) has a heavy tail; as a result, there is a subset of agents in the voting population that drives variation in the average local unemployment rate due to their relatively large influence. Theorems 1.15 and 1.16 in Section 1.6 precisely show how these two features of the composite graph generate such a strongly non-degenerate distribution, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}\left(t\right)$, even for large N.

The economy here exhibits sufficiently large configuration dependence that its outcome, determined by the outcome of the U.S. presidential election, depends on the actual allocation of unemployment among voters:

Example 1.3 (Voting Outcome, Composite Graph) *Aggregate voting behavior is characterized by Equation 1.4. The unemployment rate is 9.6 percent. Given that voters' observation network is the composite graph, and voters equally weight each of their observations, there is a probability of 0.0707 that Trump's expected vote share exceeds 0.5:*

$$\Pr\left[\frac{1}{N}\sum_{i=1}^{N}\pi_{iT} > 0.5\right] = \Pr\left[\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right) > 0.10\right] \approx 0.0707.$$

In Section 1.6, we show how to compute this quantity by hand. If we only observe the global 9.6-percent unemployment rate, we might think that the election outcome favors Clinton with certainty. However, since the average local unemployment rate can meaningfully deviate from 9.6 percent, the outcome can favor Trump, thereby setting the economy along a path that differs from the one in which Clinton is elected. The probability that the election outcome favors Clinton is 92.93 percent, and the probability that the election outcome favors Trump is 7.07 percent.

Appendix A.2 considers the distribution of possible outcomes for the economy for two variants of the composite graph. For both variants, the composite graph is constructed by pooling linkages from the base graph and the media graph, as is done here. In the first variant, the base graph is modified while the media graph stays the same. The base graph is constructed by assuming that agents form an average of 20 reciprocal linkages with other voters, rather than 50 reciprocal linkages; every voter has a self-loop as well. Given this composite graph and a 9.6-percent unemployment rate, the minimum possible average local unemployment rate is 3.63 percent, the maximum possible average local unemployment rate is 47.6 percent, and the standard deviation of the distribution of average local unemployment rates is 0.433 percentage points, so that a two-standard-deviation band equals 1.73 percentage points. The probability that the election outcome favors Trump is equal to 17.7 percent, and the probability that the election outcome favors Clinton is equal to 82.3 percent. In the second variant, the base graph stays the same, that is, agents each have a self-loop and form 50 reciprocal linkages with other voters on average, while the media graph is modified. The media graph in this second variant is constructed by assuming that each news/talk media source publishes five stories, rather than 15 stories, about the issue of jobs and unemployment. Given this composite graph and a 9.6-percent unemployment rate, the minimum possible average local unemployment rate is 6.55 percent, the maximum possible average local unemployment rate is 21.5 percent, and the standard deviation of the distribution of average local unemployment rates is 0.205 percentage points, so that a two-standard-deviation band equals 0.820 percentage points. The probability that the election outcome favors Trump is equal to 3.27 percent, and the probability that the election outcome favors Clinton is equal to 96.73 percent. For all of the graphs considered in this work, there is inherent randomness; the base graph is constructed by drawing linkages randomly between pairs of individuals, and the media graph is constructed by randomly selecting, for each news/talk media source, audience members and featured individuals. Such randomness serves as a natural benchmark; we would need to separately explore whether deviations from randomness systematically change the properties of the distribution of possible average local unemployment rates.

Each individual's local unemployment rate essentially serves as a proxy for his or her sentiment about the macroeconomy. The average local unemployment rate is thus an aggregate statistic summarizing sentiment for all agents in the system. Such sentiment, and its fluctuation, is a manifestation of animal spirits at its core. Holding the fundamentals of the economy fixed, there can be configuration-induced variations in sentiment. There can be waves of optimism if the average local unemployment rate is less than its global value, and there can be waves of pessimism if the average local unemployment rate is greater than its global value. Sentiment is moreover quantifiable; the extent to which aggregate sentiment deviates from a level that is commensurate with fundamentals depends on the extent to which the average local unemployment rate deviates from the global rate of unemployment. The underlying interaction structure among agents in the economy shapes the capacity for there to exist non-fundamental swings in aggregate sentiment. This work therefore provides a microfoundation for animal spirits.

This work buttresses other research that studies sentiment and consumer confidence, and shocks to sentiment and consumer confidence, in the macroeconomic setting and their effects on business cycles and aggregate fluctuations: for example, Farmer and Guo (1994), Barsky and Sims (2012), Angeletos and La'O (2013), Benhabib, Wang, and Wen (2015), Huo and Takayama (2015), Acharya, Benhabib, and Huo (2017), Angeletos, Collard, and Dellas (2017), and Milani (2017). In this work, we provide a simple mechanism for generating fluctuations in sentiment. Fluctuations in sentiment or animal spirits arise here from variations in configuration, with the scope for such fluctuation dependent on the topology of agents' interaction network. Cross-sectionally, variations in agents' sentiment arise from differences in agents' local environments due to differences in network position, holding the economy's fundamentals fixed.

1.3.4 When Configurations are Not Equally Likely: Distribution of Possible Average Local Unemployment Rates and the Expected Voting Outcome

Thus far, we have been considering the case in which each configuration of unemployment is equally likely to occur, meaning that each individual in the population is equally likely to be unemployed. We proceed to dispense with this assumption and instead compute the mean and variance of the distribution of possible average local unemployment rates when configurations are no longer equally likely. Now, the mean average local unemployment rate can deviate quite strongly away from the actual global rate of unemployment, f.

We segment the population into two groups: (1) those agents featured by news/talk media and (2) those agents not featured by news/talk media, with agents in the first group relatively more likely to be unemployed. The number of agents in the first group is $x = 28,003$, and the number of agents in the second group is $N - x = 137.5 \times 10^6 - 28,003$. We re-index the population of agents so that those in group 1 have indices 1 to x while those in group 2 have indices $x + 1$ to N. Attribute γ_i characterizes agents according to whether or not they have been featured by news/talk media. The probability ϕ_i that agent i is unemployed is $\phi_i = \rho_1$ for all

$i \in \{1, \ldots, x\}$ and $\phi_i = \rho_2$ for all $i \in \{x+1, \ldots, N\}$. The odds ratio for agents in group 1 relative to group 2 is: $\widehat{\psi}_1 = \frac{\frac{\rho_1}{1-\rho_1}}{\frac{\rho_2}{1-\rho_2}}$.

Example 1.4 (Configurations Unequally Likely, $\widehat{\psi}_1 = 9.42$) *Suppose that media outlets engage in "fair and balanced" reporting, providing equal air time (or equal space for hard-copy publications) to those agents who are employed and unemployed. Setting* $\rho_1 = 0.50$ *and* $\rho_2 = 0.096$, $E\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = 0.194$ *and* $\text{Std. Dev.}\, \widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = 0.00452$.

We show how to compute these first two moments in Section 1.7, employing Theorem 1.17. Here we observe extreme bias in the distribution of average local unemployment rates relative to the actual global unemployment rate of 9.6 percent. The mean average local unemployment rate is very high at 19.4 percent, so the voting outcome overwhelmingly favors Trump.

This heightened exposure to unemployment via the media might be a reason why residents of the United States, as well as residents of other countries, grossly overestimate the national unemployment rate. In an August 2014 Ipsos-MORI poll, surveyed Americans stated, on average, that the national unemployment rate was 32 percent, greatly exceeding even the U-6 unemployment rate of 11.9 percent.[14] Similarly, in an October 19-22, 2016 survey of 1000 unemployed American adults, about one in three individuals believed that the national unemployment rate was 15 percent or higher.[15] The polled individuals were unemployed, so it makes sense that they sense a national unemployment rate that exceeds the actual one. However, the extent of miscalculation is still quite significant, for in some cases, they were even provided with information about the national unemployment rate. The

[14] "Americans think the unemployment rate is 32 percent," *Vox*, November 15, 2014.

[15] "Trump supporters vastly overestimate unemployment– and they blame politicians for it," *The Washington Post*, November 2, 2016.

unemployment rate is one of the most salient features of an economy, so individuals' local perceptions of this statistic directly affect how they perceive the economy's overall health. Persistently high assessments of the unemployment rate can impact the behavior of various agents in the economy, whether such decision-making concerns voting or something else.

1.4 Sample Network-Derived Vectors of Agent Weights

Beginning with this section, we develop the mathematics that enables us to first construct the precursor distribution of possible local relative frequencies of the attribute and then construct the distribution of possible outcomes for the economy given the topology of agents' interaction network. To construct the precursor distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$, we decompose each quantity $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ into a network-derived vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$, and a configuration vector: $\mathbf{b}(N, n)$:

$$x(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N, n).$$

The vector of agent weights is the object by which the topology of agents' interaction network shapes the precursor distribution and the distribution of possible outcomes for the economy.

Quantities $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and $\mathbf{w}(\bar{\mathbf{A}})$ are both very useful objects that will allow us to study and characterize the precursor distribution and the distribution of possible outcomes for the economy, but ultimately they are abstract. We refer to $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ as the local relative frequency of the attribute; however, without additional details about the specific interpretation of this object, there does not actually exist a method by which we compute this quantity. Similarly, $\mathbf{w}(\bar{\mathbf{A}})$ is a network-derived vector of agent weights, but without additional details about this object, we do not have a method for deriving this vector. In the present section, we study different cases of $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and $\mathbf{w}(\bar{\mathbf{A}})$. For each case, we have enough detail that we can assign a particular interpretation to $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, and we can show how the vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$, naturally emerges from the underlying network. In the previous two sections, we focused on the pair $(x(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{w}(\bar{\mathbf{A}})) = \left(\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{d}_{\bar{w}}^{-}(\bar{\mathbf{A}})\right)$; in this section, we introduce additional pairs that can be relevant for other applications. Our sample set of scalar quantities and the vectors of agent weights to which they pair is as follows:

1. $\left(\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{w}_{a,i}(\bar{\mathbf{A}})\right)$. For every agent $i \in \{1, \ldots, N\}$, $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific weighted local relative frequency of the attribute in agent i's immediate network neighborhood. The accompanying random variable is $\widehat{F}_i(\bar{\mathbf{A}}, N, n)$ with CDF $G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(t)$. The corresponding network-derived vector of agent weights is $\mathbf{w}_{a,i}(\bar{\mathbf{A}}) = ([\bar{\mathbf{A}}]_{i*})^T$, so $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{w}_{a,i}(\bar{\mathbf{A}})]^T \mathbf{b}(N, n)$.

2. $\left(\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right)$. For every agent $i \in \{1, \ldots, N\}$, $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific weighted local relative frequency of the attribute for agent i following q rounds of repeated linear updating by each agent with his or her immediate neighbors. The accompanying random variable is $\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)$ with CDF $G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}},N,n)}(t)$. The corresponding network-derived vector of agent weights is $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}}) = ([\bar{\mathbf{A}}^q]_{i*})^T$, so $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n)$.

3. $\left(\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{d}_{\bar{w}}^{-}(\bar{\mathbf{A}})\right)$. $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific population-averaged weighted local relative frequency of the attribute in each agent's immediate neighborhood. The accompanying random variable is $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ with CDF $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$. The corresponding network-derived vector of agent weights is $\mathbf{d}_{\bar{w}}^{-}(\bar{\mathbf{A}}) = \frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1}$, so $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{d}_{\bar{w}}^{-}(\bar{\mathbf{A}})]^T \mathbf{b}(N, n)$.

4. $\left(\widehat{f}^{(q)}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right), \mathbf{d}_w^{-(q)}\left(\bar{\mathbf{A}}\right)\right)$. $\widehat{f}^{(q)}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is the configuration-specific population-averaged weighted local relative frequency of the attribute following q rounds of repeated linear updating by each agent with his or her immediate neighbors. The accompanying random variable is $\widehat{F}^{(q)}_{avg}\left(\bar{\mathbf{A}}, N, n\right)$ with CDF $G_{\widehat{F}^{(q)}_{avg}(\bar{\mathbf{A}},N,n)}(t)$. The corresponding network-derived vector of agent weights is $\mathbf{d}_w^{-(q)}\left(\bar{\mathbf{A}}\right) = \frac{1}{N}\left[\bar{\mathbf{A}}^q\right]^T \mathbf{1}$, so $\widehat{f}^{(q)}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{d}_w^{-(q)}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right)$.

5. $\left(\widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right), \mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right)$. $\widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is the configuration-specific consensus local relative frequency of the attribute following infinitely many rounds of repeated linear updating by each agent with his or her immediate neighbors. The accompanying random variable is $\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n\right)$ with CDF $G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(t)$. The corresponding network-derived vector of agent weights is $\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)$, to be later computed in this section, and $\widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right)$.

Observe that each of these scalar quantities, that is, each of these local relative frequencies of the attribute, is computed in a manner exactly parallel to the way that we computed $x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$; for example:

$$x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right) \text{ and}$$
$$\widehat{f}^{(q)}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{d}_w^{-(q)}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right).$$

Also note that the computation of each CDF parallels the computation of $G_{X(\bar{\mathbf{A}},N,n)}(t)$. We additionally observe that each of the vectors of agent weights derived above has elements that sum to 1.[16]

We would like to now demonstrate how to compute $\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)$, and we would like to show how the elements of this vector depend on network primitives when

[16]To see this, note that $\bar{\mathbf{A}}$ is row-stochastic, and row-stochasticity is preserved under matrix multiplication.

graph $\mathcal{G}(\mathbf{A})$ satisfies particular assumptions. First, we define the period of a node and aperiodicity of a graph:

Definition 1.2 *For a graph $\mathcal{G}(\mathbf{A}) = (\mathcal{V}(\mathbf{A}), \mathcal{E}(\mathbf{A}))$, the* period *$pd_i$ of node $i \in \mathcal{V}(\mathbf{A})$ is $pd_i \equiv gcd\{q \geq 1 : [\mathbf{A}^q]_{ii} > 0\}$, where gcd denotes the greatest common divisor. Node i is* aperiodic *when $pd_i = 1$, and graph $\mathcal{G}(\mathbf{A}) = (\mathcal{V}(\mathbf{A}), \mathcal{E}(\mathbf{A}))$ is* aperiodic *when $pd_i = 1$ for all nodes $i \in \mathcal{V}(\mathbf{A})$.*

To compute the period of a node on a graph, construct a set that contains the lengths of all possible cycles for that node and then identify the greatest common divisor among all integers in that set. A node is aperiodic when the greatest common divisor among all integers in that set equals 1, and a graph is aperiodic when every constituent node is itself aperiodic.

To compute $\mathbf{w}_\infty(\bar{\mathbf{A}})$, one necessary assumption is that the row-stochastic weighted adjacency matrix $\bar{\mathbf{A}}$ must be primitive. Primitivity of $\bar{\mathbf{A}}$ is equivalent to strong connectedness and aperiodicity of its directed companion graph $\mathcal{G}(\bar{\mathbf{A}})$. Defining $w_{ij}^{(q)} \equiv [\bar{\mathbf{A}}^q]_{ij}$ as the weight that agent i assigns to agent j following q rounds of linear updating, we now demonstrate the existence of $\mathbf{w}_\infty(\bar{\mathbf{A}})$ and its computation:

Theorem 1.1 *If $\bar{\mathbf{A}}$ is primitive, then $\lim_{q\to\infty} w_{ij}^{(q)} = [\mathbf{w}_\infty^T]_j$ exists. The pair $(\mathbf{w}_\infty^T, 1)$ is the unique dominant left eigenpair of $\bar{\mathbf{A}}$, $\mathbf{w}_\infty^T \bar{\mathbf{A}} = \mathbf{w}_\infty^T$, and $\mathbf{w}_\infty^T \mathbf{1} = 1$.*

Provided that $\bar{\mathbf{A}}$ is primitive, we compute vector $\mathbf{w}_\infty(\bar{\mathbf{A}})$ by solving for the left eigenvector of $\bar{\mathbf{A}}$ corresponding to its unit eigenvalue, which happens to be the dominant, or largest, eigenvalue for $\bar{\mathbf{A}}$. As $q \to \infty$, provided that $\bar{\mathbf{A}}$ is primitive, the weight that every agent assigns to agent j converges to the same limiting value,

$[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j$. Specifically,

$$\lim_{q\to\infty} \bar{\mathbf{A}}^q = \begin{pmatrix} \text{—} & \mathbf{w}_\infty^T & \text{—} \\ & \vdots & \\ \text{—} & \mathbf{w}_\infty^T & \text{—} \end{pmatrix}.$$

Since every agent assigns the same weight to each agent j in the population, in this setting, the local relative frequency of the attribute for any configuration, $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{w}_\infty(\bar{\mathbf{A}})]^T \mathbf{b}(N, n)$, is the same across agents. There is consensus among agents, and we refer to $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ as the consensus local relative frequency of the attribute. A variant of Theorem 1.1 is presented in DeGroot (1974), with infinite repeated linear updating and the vector, $\mathbf{w}_\infty(\bar{\mathbf{A}})$, of agent weights under consensus forming the basis for DeGroot learning.

Vector $\mathbf{w}_\infty(\bar{\mathbf{A}})$ can be expressed in a closed form provided that $\mathcal{G}(\mathbf{A})$, the unweighted graph that pairs with graph $\mathcal{G}(\bar{\mathbf{A}})$, satisfies certain assumptions:

Theorem 1.2 *If graph $\mathcal{G}(\mathbf{A}) = (\mathcal{V}(\mathbf{A}), \mathcal{E}(\mathbf{A}))$ is undirected, connected, and aperiodic, and all non-zero elements within every row of the corresponding matrix $\bar{\mathbf{A}}$ have the same value, then $\mathbf{w}_\infty(\bar{\mathbf{A}}) = \frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}} > \mathbf{0}$, where $[\mathbf{d}]_i$ is the degree for agent i.*

Provided that we satisfy the assumptions of Theorem 1.2, an agent's weight under consensus is directly proportional to his or her degree: $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_i = \frac{d_i}{\mathbf{1}^T\mathbf{d}}$. A self-loop increases an agent's degree by one unit, which makes the degree vector equal to: $\mathbf{d} = \mathbf{A}\mathbf{1}$. A result similar to that of Theorem 1.2 appears in DeMarzo, Vayanos, and Zwiebel (Theorem 6, 2003).

We can also establish a closed-form solution for $\mathbf{w}_\infty(\bar{\mathbf{A}})$ when $\mathcal{G}(\mathbf{A})$ is directed, aperiodic, and Eulerian. We define a directed graph to be *Eulerian* when it has the following properties:

Definition 1.3 *A directed graph* $\mathcal{G}(\mathbf{A}) = (\mathcal{V}(\mathbf{A}), \mathcal{E}(\mathbf{A}))$ *is* Eulerian *if and only if it is strongly connected and* $\mathbf{d}^{+} = \mathbf{d}^{-}$.

A directed graph is Eulerian when each node's in-degree equals its out-degree, and the graph is strongly connected. We can now present the closed-form solution for $\mathbf{w}_{\infty}(\bar{\mathbf{A}})$:

Theorem 1.3 *If the graph* $\mathcal{G}(\mathbf{A}) = (\mathcal{V}(\mathbf{A}), \mathcal{E}(\mathbf{A}))$ *is Eulerian and aperiodic, and all non-zero elements within every row of the corresponding matrix* $\bar{\mathbf{A}}$ *have the same value, then* $\mathbf{w}_{\infty}(\bar{\mathbf{A}}) = \frac{\mathbf{d}^{+}}{\mathbf{1}^{T}\mathbf{d}^{+}} > \mathbf{0}$, *where* $[\mathbf{d}^{+}]_{i}$ *is the out-degree for agent i.*

For this class of graphs, an agent's weight under consensus is directly proportional to his or her out-degree, or equivalently, his or her in-degree: $[\mathbf{w}_{\infty}(\bar{\mathbf{A}})]_{i} = \frac{d_{i}^{+}}{\mathbf{1}^{T}\mathbf{d}^{+}} = \frac{d_{i}^{-}}{\mathbf{1}^{T}\mathbf{d}^{-}}$. A self-loop increases both an agent's out-degree and in-degree by one unit, so $\mathbf{d}^{+} = \mathbf{d}^{-} = \mathbf{A}\mathbf{1} = \mathbf{A}^{T}\mathbf{1}$.

We can additionally establish a closed-form solution for $\mathbf{w}_{\infty}(\bar{\mathbf{A}})$ for the family of random digraphs with N nodes, no self-loops, probability p of directed edge formation independent across edges, and symmetric edge weights so that $[\bar{\mathbf{A}}]_{ij} = \frac{1}{d_{i}^{+}}$ if there exists a directed edge from node i to node j:

Theorem 1.4 *If* $(\chi(N) - 1)\log N \to \infty$, *where* $Np = \chi(N)\log N$, *then w.h.p.* $\mathbf{w}_{\infty}(\bar{\mathbf{A}})$ *is unique and* $\mathbf{w}_{\infty}(\bar{\mathbf{A}}) \sim \frac{\mathbf{d}^{-}+\boldsymbol{\iota}}{E[|\mathcal{E}|]}$, *where* $[\boldsymbol{\iota}]_{i} = \max_{j\in\mathcal{N}^{-}(i)} \frac{d_{j}^{-}}{d_{j}^{+}}$ *and* $\mathcal{N}^{-}(i)$ *is the in-neighborhood of node i. w.h.p.* $\mathbf{w}_{\infty}(\bar{\mathbf{A}}) \sim \frac{\mathbf{d}^{-}}{E[|\mathcal{E}|]}$ *for* $N - o\left(N^{1/4}\right)$ *nodes. If* $\chi(N) = 1 + \kappa$, $\kappa > 0$, *or* $(\chi(N) - 1)\log N = \omega(\log\log N)$, *then w.h.p.* $\mathbf{w}_{\infty}(\bar{\mathbf{A}}) \sim \frac{\mathbf{d}^{-}}{E[|\mathcal{E}|]}$.

The statements of this theorem are made w.h.p. relative to the family of random digraphs. We take $(\chi(N) - 1)\log N \to \infty$ so that each random digraph is strongly connected w.h.p. Given that we are considering random digraphs with no self-loops, the expected number of edges for each graph is $E[|\mathcal{E}|] = N(N-1)p \sim N^{2}p$. As

ι_i increases, one of agent i's in-neighbors becomes relatively more observable, so $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_i$ consequently increases. In general, the higher an agent's in-degree, the higher that agent's weight. This theorem can be formulated due to parallels between the vector $\mathbf{w}_\infty(\bar{\mathbf{A}})$ for a random digraph with symmetrically weighted edges and the stationary distribution for a simple random walk on a random digraph. Cooper and Frieze (2012) study this latter problem, so we can therefore adapt some of their mathematics to provide insight into the behavior of $\mathbf{w}_\infty(\bar{\mathbf{A}})$ over this class of random digraphs.

These newly introduced closed-form expressions for $\mathbf{w}_\infty(\bar{\mathbf{A}})$, in addition to serving as sample network-derived vectors of agent weights, can also be used in research on DeGroot learning. The next example computes the set of network-derived vectors of agent weights introduced at the beginning of this section for a network with 15 nodes formed from preferential attachment:

Example 1.5 (Network-Derived Vectors of Agent Weights) *Consider an economy with $N = 15$ agents whose interaction structure $\mathcal{G}(\mathbf{A})$ is depicted in the top left of Figure 1.9. Assume that agents equally weight each of their linkages. Figure 1.9 plots the following vectors of agent weights for $i = 1$ and $q = 5$: $\mathbf{w}_{a,i}(\bar{\mathbf{A}})$, $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})$, $\mathbf{d}_w^-(\bar{\mathbf{A}})$, $\mathbf{d}_w^{-(q)}(\bar{\mathbf{A}})$, and $\mathbf{w}_\infty(\bar{\mathbf{A}})$.*

Since graph $\mathcal{G}(\mathbf{A})$ in Example 1.5 is undirected, connected, and aperiodic, and agents equally weight each of their linkages, by Theorem 1.2, each agent's weight under consensus is directly proportional to his or her degree. We observe this relation between an agent's degree and his or her weight under consensus in the bottom right plot of Figure 1.9. With $\lim_{q\to\infty}\bar{\mathbf{A}}^q = \mathbf{1}[\mathbf{w}_\infty(\bar{\mathbf{A}})]^T$ and convergence of $\bar{\mathbf{A}}^q$ fast in this particular setting, $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}}) \approx \mathbf{w}_\infty(\bar{\mathbf{A}})$ and $\mathbf{d}_w^{-(q)}(\bar{\mathbf{A}}) \approx \mathbf{w}_\infty(\bar{\mathbf{A}})$, which we also observe in Figure 1.9.

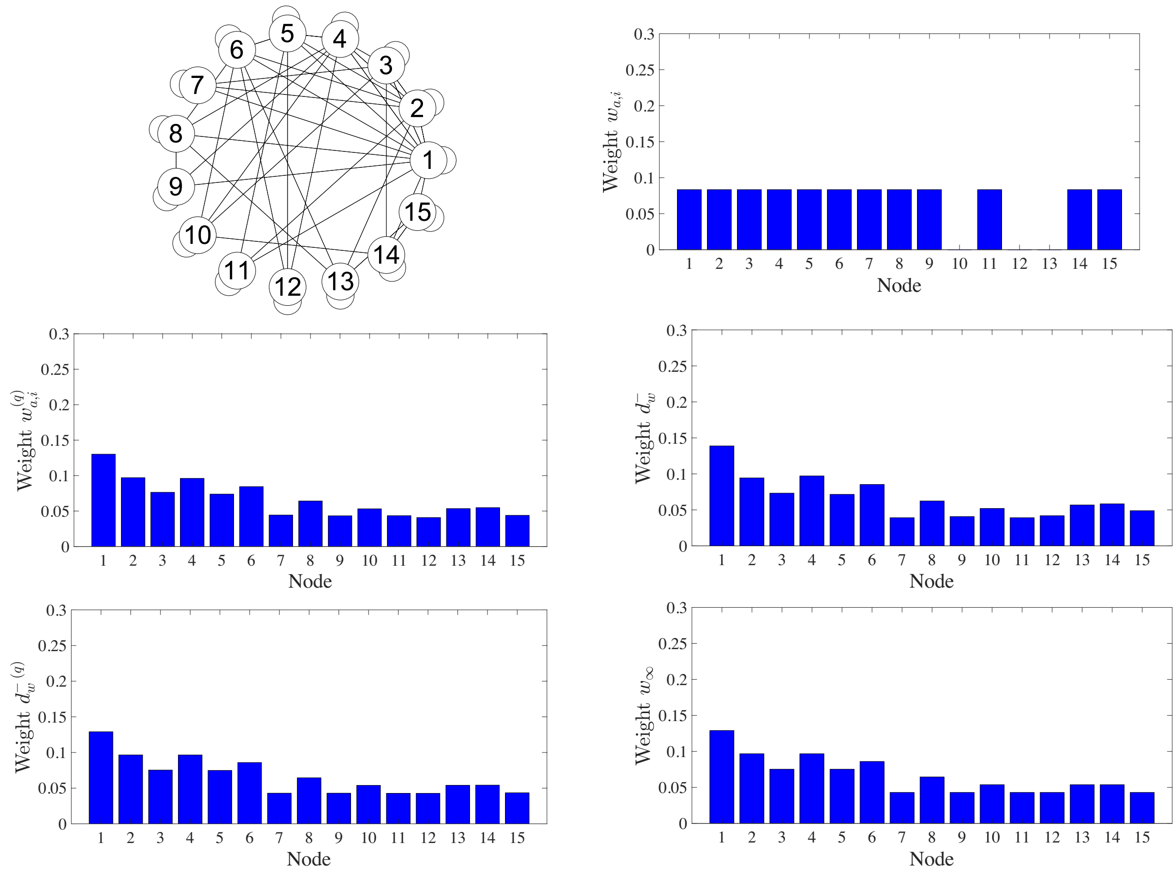


Figure 1.9: *Graph $\mathcal{G}(\mathbf{A})$ from Example 1.5 and plots of $\mathbf{w}_{a,i}(\bar{\mathbf{A}})$, $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})$, $\mathbf{d}_w^-(\bar{\mathbf{A}})$, $\mathbf{d}_w^{-(q)}(\bar{\mathbf{A}})$, and $\mathbf{w}_\infty(\bar{\mathbf{A}})$, setting $i = 1$ and $q = 5$, and assuming that agents equally weight their linkages.*

Appendix A.3 characterizes $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}}) = \left([\bar{\mathbf{A}}^q]_{i*}\right)^T$ and $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n)$ for all finite q and in the limit as $q \to \infty$ in terms of the fundamental features of $\bar{\mathbf{A}}$. It also studies the rate of convergence of $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$.

1.5 When Configuration is Irrelevant: The Degeneracy of $G_{X(\bar{\mathbf{A}},N,n)}(t)$

In the previous section, we studied particular cases of $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and $\mathbf{w}(\bar{\mathbf{A}})$. For each case, we showed how to derive the vector of agent weights from the underlying network, and we examined how agents' weights depend on the topological features of the network. This list of weighting vectors is certainly not exhaustive; the relevant vector of agent weights, in general, emerges naturally from the setting under study.

We now focus on the relationship between the network-derived vector of agent weights and the precursor distribution of possible local relative frequencies of the attribute. For this section, we first study the null case in which the probability distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$, is degenerate, meaning that the support of $X(\bar{\mathbf{A}}, N, n)$ is uniquely valued. We characterize the necessary and sufficient restrictions on both the vector of agent weights and the underlying network structure for the precursor distribution to be degenerate. Degeneracy of this probability distribution can lead to the distribution of possible outcomes for the economy also being degenerate. After studying the case in which the probability distribution is degenerate, we then determine the necessary and sufficient conditions for the population-wide cross-sectional distribution of local relative frequencies of the attribute to be invariant to configuration. In such a setting, the population set of actions becomes invariant to configuration, which makes the outcome of the economy also invariant to configuration and the distribution of possible outcomes degenerate. We find that the conditions for degeneracy are quite restrictive. Most economic systems with interacting agents therefore tend to have probability distributions that feature some level of non-degeneracy, which makes the outcome of the economy dependent on the particular configuration of the attribute among agents. Section 1.6 studies the more general case of non-degeneracy.

We begin by focusing on the precursor distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$ of possible local relative frequencies of the attribute. Degeneracy of this distribution arises when quantity $x(\bar{\mathbf{A}},\mathbf{b},N,n)$ is fixed for all possible configurations $\mathbf{b} \in \mathcal{B}(N,n)$. When this condition holds for every feasible f, we say that quantity $x(\bar{\mathbf{A}},\mathbf{b},N,n)$ is *invariant to configuration*:

Definition 1.4 *Quantity* $x(\bar{\mathbf{A}},\mathbf{b},N,n)$ *is* invariant to configuration *when* $x(\bar{\mathbf{A}},\mathbf{b},N,n) = x(\bar{\mathbf{A}},\mathbf{b}',N,n)$ *for all configurations* $\mathbf{b},\mathbf{b}' \in \mathcal{B}(N,n)$*, and this property holds for all feasible* n*.*

When $x(\bar{\mathbf{A}},\mathbf{b},N,n)$ is invariant to configuration for each feasible level f, the support of $X(\bar{\mathbf{A}},N,n)$ takes one value and the variance of $X(\bar{\mathbf{A}},N,n)$ is zero. In the next theorem, we determine the necessary and sufficient restrictions on the corresponding vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$, for $x(\bar{\mathbf{A}},\mathbf{b},N,n)$ to be invariant to configuration, and we solve for the support of $X(\bar{\mathbf{A}},N,n)$:

Theorem 1.5 *Scalar quantity* $x(\bar{\mathbf{A}},\mathbf{b},N,n) = [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N,n)$ *is invariant to configuration if and only if* $[\mathbf{w}(\bar{\mathbf{A}})]_i = \frac{1}{N}$ *for all* $i \in \{1,\ldots,N\}$*. When* $x(\bar{\mathbf{A}},\mathbf{b},N,n)$ *is invariant to configuration,* $x(\bar{\mathbf{A}},\mathbf{b},N,n) = \frac{n}{N}$*.*

When $\mathbf{w}(\bar{\mathbf{A}}) = \frac{1}{N}\mathbf{1}$, every agent has the same effective representation in the population. As a result, regardless of which agents have the attribute's unit value, the overall contribution by those agents to that attribute's local relative frequency is the same. The local relative frequency of the attribute is always equal to the attribute's global relative frequency, f. The support of $X(\bar{\mathbf{A}},N,n)$ is thus f. The configuration of the attribute among agents in the system is irrelevant for how the system evolves, and the outcome of the economy only depends on the system's aggregate feature, f.

We can similarly establish the necessary and sufficient restrictions on specific network-derived vectors of agent weights for their corresponding local relative frequencies of the binary-valued attribute to be invariant to configuration:

Corollary 1.1 (to Theorem 1.5) (1) *Scalars* $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *and* $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *are respectively invariant to configuration if and only if, for all* $j \in \{1, \dots, N\}$, $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_j = \frac{1}{N}$ *and* $\left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]_j = \frac{1}{N}$.
(2) *Vectors* $\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *and* $\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *are respectively invariant to configuration if and only if, for all* $i, j \in \{1, \dots, N\}$, $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_j = \frac{1}{N}$ *and* $\left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]_j = \frac{1}{N}$.
(3) *Scalars* $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *and* $\widehat{f}_{avg}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *are respectively invariant to configuration if and only if, for all* $i \in \{1, \dots, N\}$, $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_i = \frac{1}{N}$ *and* $\left[\mathbf{d}_w^{-(q)}(\bar{\mathbf{A}})\right]_i = \frac{1}{N}$.
(4) *Scalar* $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *is invariant to configuration if and only if, for all* $i \in \{1, \dots, N\}$, $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_i = \frac{1}{N}$.

When $x(\bar{\mathbf{A}}, \mathbf{b}, N, n) \in \left\{ \left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^N, \left\{\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^N, \widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \widehat{f}_{avg}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \right\}$ *is invariant to configuration,* $x(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \frac{n}{N}$. *When* $\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *and* $\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ *are invariant to configuration,* $\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \frac{n}{N}\mathbf{1}$.

Corollary 1.1 follows immediately from Theorem 1.5. The scalar quantities, $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, $\widehat{f}_{avg}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, and $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ are all invariant to configuration when the vectors of agent weights to which they are paired, equal $\frac{1}{N}\mathbf{1}$. Then, every agent has the same effective representation in the population, so regardless of which subset of agents has the attribute's unit value, the local relative frequency of that attribute remains the same. The distributions $G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(t)$, $G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}},N,n)}(t)$, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, $G_{\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}},N,n)}(t)$, and $G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(t)$ all become degenerate; their supports are uniquely valued and equal to the attribute's global relative frequency, f. The local relative frequency of the attribute therefore equals its global relative frequency with probability 1, and this can make the distribution of possible outcomes for the economy degenerate and only dependent on the system's aggregate feature, f.

Returning to Corollary 1.1, vectors $\widehat{\mathbf{f}}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ and $\widehat{\mathbf{f}}^{(q)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ are invariant to configuration when each of their respective elements, $\widehat{f}_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ and $\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$, are invariant to configuration. When we have invariance along each dimension, both the population vector of agent actions and the population set of agent actions are invariant to configuration provided that each agent chooses an action that depends on his or her local relative frequency of the attribute. The outcome of the economy then becomes unique. When we have invariance along each dimension, we can also define multivariate distributions, $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}\left(\mathbf{t}\right)$ and $G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}},N,n)}\left(\mathbf{t}\right)$, whose N marginal distributions are all degenerate.

In Corollary 1.1, we stated the necessary and sufficient restrictions on specific vectors of agent weights that lead to degenerate probability distributions. Since these vectors of agent weights are explicitly network-derived, Theorem 1.6 presents the corresponding necessary and sufficient restrictions on $\bar{\mathbf{A}}$:

Theorem 1.6 (1) *Scalars* $\widehat{f}_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *and* $\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *are respectively invariant to configuration if and only if* $[\bar{\mathbf{A}}]_{i*}=\frac{1}{N}\mathbf{1}^T$ *and* $[\bar{\mathbf{A}}^q]_{i*}=\frac{1}{N}\mathbf{1}^T$. (2) *Vectors* $\widehat{\mathbf{f}}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *and* $\widehat{\mathbf{f}}^{(q)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *are respectively invariant to configuration if and only if* $\bar{\mathbf{A}}=\frac{1}{N}\mathbf{1}\mathbf{1}^T$ *and* $\bar{\mathbf{A}}^q=\frac{1}{N}\mathbf{1}\mathbf{1}^T$. (3) *Scalars* $\widehat{f}_{avg}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *and* $\widehat{f}_{avg}^{(q)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *are respectively invariant to configuration if and only if* $\bar{\mathbf{A}}$ *is doubly stochastic and* $\bar{\mathbf{A}}^q$ *is doubly stochastic.* (4) *Scalar* $\widehat{f}^{(\infty)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)$ *is invariant to configuration if and only if* $\bar{\mathbf{A}}$ *is doubly stochastic.*

When these restrictions on $\bar{\mathbf{A}}$ or $\bar{\mathbf{A}}^q$ are satisfied, the distributions $G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}\left(t\right)$, $G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}},N,n)}\left(t\right)$, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}\left(t\right)$, $G_{\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}},N,n)}\left(t\right)$, and $G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}\left(t\right)$ are all degenerate for every feasible f.

Let us first discern the relationship between the double stochasticity of $\bar{\mathbf{A}}$ ($\bar{\mathbf{A}}^q$) and the invariance of $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ ($\widehat{f}_{avg}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to configuration. The sum of the j^{th} column of $\bar{\mathbf{A}}$ ($\bar{\mathbf{A}}^q$) represents the total weight that every agent in the population accords to agent j; it is the effective representation of agent j in the population. When $\bar{\mathbf{A}}$ ($\bar{\mathbf{A}}^q$) is doubly stochastic, every agent in the population has the same effective representation, that of one agent, so configuration becomes irrelevant. We now discern the relationship between the double stochasticity of $\bar{\mathbf{A}}$ and the invariance of $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to configuration. When $\bar{\mathbf{A}}$ is both primitive and doubly stochastic, $\lim_{q\to\infty} \bar{\mathbf{A}}^q = \mathbf{1}\left[\mathbf{w}_\infty(\bar{\mathbf{A}})\right]^T$ is also doubly stochastic. Since $\left[\mathbf{1}\mathbf{w}_\infty^T(\bar{\mathbf{A}})\right]_{ij} = \left[\mathbf{w}_\infty(\bar{\mathbf{A}})\right]_j$ for each $i \in \{1, \ldots, N\}$, $\lim_{q\to\infty} \bar{\mathbf{A}}^q = \frac{1}{N}\mathbf{1}\mathbf{1}^T$ with $\mathbf{w}_\infty(\bar{\mathbf{A}}) = \frac{1}{N}\mathbf{1}$; every agent has the same weight under consensus when $\bar{\mathbf{A}}$ is doubly stochastic, which makes $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ invariant to configuration. The necessary restrictions on $\bar{\mathbf{A}}$ for these various local relative frequencies of the attribute to be exactly invariant to configuration are very limiting. Many underlying graphs thus generate some level of non-degeneracy and dependence on configuration.

In the next example, we illustrate how $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ becomes invariant to configuration and $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ becomes degenerate for all f when $\bar{\mathbf{A}}$ is doubly stochastic:

Example 1.6 (Invariance of $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to Configuration) *Consider an economy with $N = 15$ agents whose interaction network $\mathcal{G}(\mathbf{A})$ (Figure 1.10) is a directed 4-regular graph with self-loops for every node. Assuming that agents equally weight each of their out-edges, $\widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \frac{n}{N}$ is invariant to configuration and distribution $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ is degenerate for all feasible n.*

When the underlying graph $\mathcal{G}(\mathbf{A})$ is regular and agents assign an equal weight to each of their linkages, $\bar{\mathbf{A}}$ becomes doubly stochastic, so $\mathbf{d}_w^-(\bar{\mathbf{A}}) = \frac{1}{N}\mathbf{1}$ (Figure 1.10, top right) and $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ does not change with configuration (Figure 1.10, bottom). For all feasible f, the population-averaged local relative frequency of the attribute equals its global relative frequency.

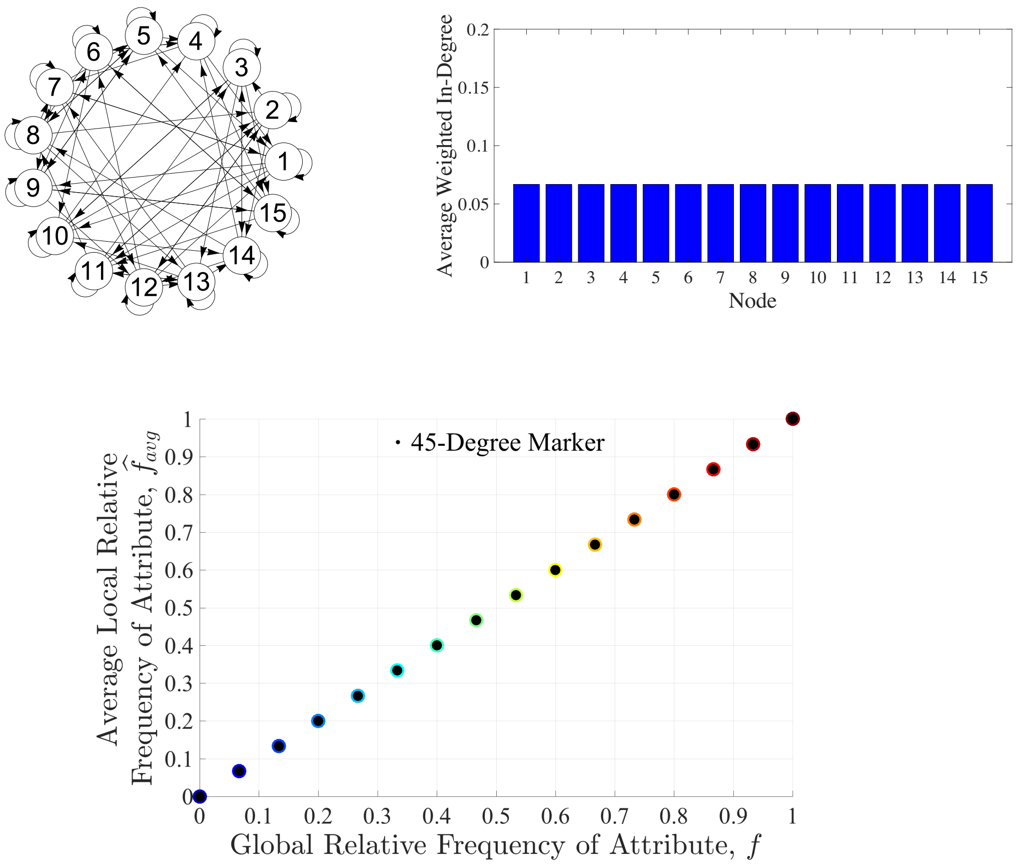


Figure 1.10: *Corresponding to Example 1.6, directed 4-regular graph with self-loops $\mathcal{G}\left(\mathbf{A}\right)$ (top left), a plot of the average weighted in-degree for each agent, $\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)$ (top right), and the average local relative frequency of the attribute, $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, for every possible configuration $\mathbf{b} \in \mathcal{B}\left(N, n\right)$ and for all feasible global relative frequencies of the attribute, f (bottom).*

We have finished examining settings in which the precursor distribution of possible local relative frequencies of the attribute is degenerate. This degeneracy of $G_{X(\bar{\mathbf{A}},N,n)}\left(t\right)$ and its counterparts is often a necessary prerequisite for the distribution of possible outcomes for the economy to also be degenerate. Degeneracy of these probability distributions means that the economy is not configuration dependent;

rather, the way that the economy evolves depends on f, the global relative frequency of the attribute and the economy's aggregate feature, and not on the particular configuration of the attribute among agents. It is very difficult to exactly satisfy the conditions that make these probability distributions degenerate, and therefore, most economic systems exhibit some level of dependence on configuration. Modeling the economy as if its evolution only depends on its aggregate features is, in general, incomplete.

We next transition towards characterizing environments in which the population-wide cross-sectional distribution of local relative frequencies of the attribute is invariant to configuration. When the cross-sectional distribution is invariant to configuration, the cross-sectional distribution of agent actions becomes invariant to configuration, which can lead to a unique outcome for the economy. The next theorem provides necessary and sufficient conditions for the unordered multi-sets, $\left\{\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$ and $\left\{\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$, to be invariant to configuration. It places necessary and sufficient restrictions on $\bar{\mathbf{A}}$ (respectively $\bar{\mathbf{A}}^q$):

Theorem 1.7 *Unordered multiset* $\left\{\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$ *(respectively unordered multiset* $\left\{\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$*) is invariant to configuration if and only if the following two conditions hold:*

(1) *the row sum of any n column vectors of* $\bar{\mathbf{A}}$ *(respectively* $\bar{\mathbf{A}}^q$*) has the same multiset of elements, and this property holds for every* $n \in \left\{1, \ldots, \lfloor \frac{N}{2} \rfloor\right\} \subseteq \mathbb{Z}_+$*, and*

(2) $\bar{\mathbf{A}}$ *(respectively* $\bar{\mathbf{A}}^q$*) is doubly stochastic.*

When these restrictions are satisfied, the cross-sectional distribution of agents' weighted local relative frequencies of the attribute is invariant to configuration for any value n. When $n = 1$, every column of $\bar{\mathbf{A}}$ (respectively $\bar{\mathbf{A}}^q$) must have the same multiset of elements. Since all matrix elements in $\bar{\mathbf{A}}$ (respectively $\bar{\mathbf{A}}^q$) must also sum to N, $\bar{\mathbf{A}}$ (respectively $\bar{\mathbf{A}}^q$) is doubly stochastic. It is very difficult to exactly satisfy the restrictions on $\bar{\mathbf{A}}$ so that the population set of weighted local relative frequencies of the attribute is invariant to configuration. Therefore, in most settings, the cross-sectional distribution of agent actions varies with configuration and the distribution of possible outcomes for the economy is non-degenerate.

Notwithstanding these strong restrictions, there do exist networks $\mathcal{G}(\bar{\mathbf{A}})$ for which the conditions of Theorem 1.7 are satisfied. When $\bar{\mathbf{A}} = \frac{1}{N}\mathbf{1}\mathbf{1}^T$, so that the underlying interaction network is a complete graph with self-loops for every node, or when $\bar{\mathbf{A}} = \mathbf{I}_{N\times N}$, so that the underlying interaction network is a graph with isolates and a self-loop for every node, multiset $\left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^{N}$ is invariant to configuration for all feasible n. In the former case, multiset $\left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^{N} = \left\{\frac{n}{N}, \frac{n}{N}, \ldots, \frac{n}{N}\right\}$, and in the latter case, multiset $\left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^{N} = \left\{\{1\}_{i=1}^{n}, \{0\}_{i=n+1}^{N}\right\}$.

1.6 When Configuration Matters: The Non-Degeneracy of $G_{X(\bar{\mathbf{A}},N,n)}(t)$

We just finished studying the null setting in which the particular configuration of the attribute among agents is irrelevant and the probability distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$ is degenerate. We now leave this null setting behind; in this section, we develop the mathematics that enables us to fully characterize the distribution of $X(\bar{\mathbf{A}}, N, n)$ when it is non-degenerate. These theoretical results hold for all possible population sizes, network topologies, and prevalences of the binary-valued attribute in the population. These theoretical findings allow us to directly map the topology of agents' interaction network to the distributional features of $X(\bar{\mathbf{A}}, N, n)$. Specifically, we map the topology of agents' interaction network, $\mathcal{G}(\bar{\mathbf{A}})$, to a network-derived vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$, and we then map the network-derived vector of agent weights to the probability distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$. The theoretical findings presented in this section allow us to collapse the complexities of network-based agent interactions into a simple probability distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$, that we can then use to construct the probability distribution of possible outcomes for the economy.

We begin by presenting results concerning the distributional features of $X(\bar{\mathbf{A}}, N, n)$, and we then explore how $G_{X(\bar{\mathbf{A}},N,n)}(t)$ can remain approximately non-degenerate even for very large N. For several findings, we assume that each configuration, $\mathbf{b}(N, n) \in \mathcal{B}(N, n)$, of the binary-valued attribute among agents in the system is equally likely.

1.6.1 Distributional Features of $X(\bar{\mathbf{A}}, N, n)$

We characterize all notable features of the distribution of possible local relative frequencies of the attribute, $G_{X(\bar{\mathbf{A}},N,n)}(t)$.[17] However, before doing so, we introduce a few more pieces of notation that simplify certain expressions later on in this section. Define random variable $W(\bar{\mathbf{A}})$ with realization $[\mathbf{w}(\bar{\mathbf{A}})]_i$, the weight for agent i. In this section, we are interested in the population moments of $W(\bar{\mathbf{A}})$. We construct these moments from the elements of the network-derived vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$; unless otherwise specified, there is no randomness in the population set of agent weights. We can similarly introduce random variables for particular cases of agent weights. Define $W_{a,j}(\bar{\mathbf{A}})$ for all $j \in \{1, \ldots, N\}$, $W_{a,j}^{(q)}(\bar{\mathbf{A}})$ for all $j \in \{1, \ldots, N\}$, $D_w^-(\bar{\mathbf{A}})$, $D_w^{-\,(q)}(\bar{\mathbf{A}})$, and $W_\infty(\bar{\mathbf{A}})$, whose realizations are respectively agent weights

[17]These results hold even when the vector of agent weights is not network-derived.

$\left[\mathbf{w}_{a,j}\left(\bar{\mathbf{A}}\right)\right]_i$, $\left[\mathbf{w}_{a,j}^{(q)}\left(\bar{\mathbf{A}}\right)\right]_i$, $\left[\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)\right]_i$, $\left[\mathbf{d}_w^{-\,(q)}\left(\bar{\mathbf{A}}\right)\right]_i$, and $\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]_i$.

We also define random variables for the degree of an undirected graph, the out-degree of a directed graph, and the in-degree of a directed graph. When $\mathcal{G}\left(\mathbf{A}\right)$ is undirected, the degree vector is $\mathbf{d}\left(\mathbf{A}\right) = \mathbf{A}\mathbf{1}$. When $\mathcal{G}\left(\mathbf{A}\right)$ is directed, the out-degree vector is $\mathbf{d}^+\left(\mathbf{A}\right) = \mathbf{A}\mathbf{1}$ and the in-degree vector is $\mathbf{d}^-\left(\mathbf{A}\right) = \mathbf{A}^T\mathbf{1}$. Define random variables $D\left(\mathbf{A}\right)$, $D^+\left(\mathbf{A}\right)$, and $D^-\left(\mathbf{A}\right)$ whose realizations are respectively the degree for agent i, $\left[\mathbf{d}\left(\mathbf{A}\right)\right]_i$, the out-degree for agent i, $\left[\mathbf{d}^+\left(\mathbf{A}\right)\right]_i$, and the in-degree for agent i, $\left[\mathbf{d}^-\left(\mathbf{A}\right)\right]_i$.

The distribution $G_{X(\bar{\mathbf{A}},N,n)}\left(t\right)$ strongly depends on the distributional features of agents' weights. We can see this relationship most clearly when $n = 1$. For that case, $G_{X(\bar{\mathbf{A}},N,n)}\left(t\right) = G_{W(\bar{\mathbf{A}})}\left(t\right)$, and the distribution of possible local relative frequencies of the attribute equals the distribution of agent weights. For the remainder of this section, we characterize the distributional features of $X\left(\bar{\mathbf{A}}, N, n\right)$ for a general n, and we determine how these features of $X\left(\bar{\mathbf{A}}, N, n\right)$ relate to network structure via the set of network-derived agent weights. We begin by defining the first moment of $X\left(\bar{\mathbf{A}}, N, n\right)$:

Theorem 1.8 $EX\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N} = f$.

The first moment of $X\left(\bar{\mathbf{A}}, N, n\right)$ is equal to the attribute's global relative frequency, f. The local relative frequency of the attribute can deviate in either direction away from the attribute's global relative frequency, but in expectation, it must equal this value. The distribution of $X\left(\bar{\mathbf{A}}, N, n\right)$ is consequently centered about the point in which configuration is irrelevant and only the aggregate feature, f, matters. Note that this result follows from assuming that each configuration is equally likely. We will see in the next section that this relationship between $EX\left(\bar{\mathbf{A}}, N, n\right)$ and f disappears once each configuration is no longer equally likely to occur.

We can extend this result about the first moment to particular cases of $X(\bar{\mathbf{A}}, N, n)$:

Corollary 1.2 (to Theorem 1.8) *For every random variable* $X(\bar{\mathbf{A}}, N, n) \in \left\{\left\{\widehat{F}_i(\bar{\mathbf{A}}, N, n)\right\}_{i=1}^{N}, \left\{\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)\right\}_{i=1}^{N}, \widehat{F}_{avg}(\bar{\mathbf{A}}, N, n), \widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}}, N, n), \widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)\right\}$, $EX(\bar{\mathbf{A}}, N, n) = \frac{n}{N}$.

For the remainder of this section, we omit the corollaries that immediately follow from each of the presented theorems. We can construct such corollaries by substituting for each theorem the triplet $(X(\bar{\mathbf{A}}, N, n), W(\bar{\mathbf{A}}), \mathbf{w}(\bar{\mathbf{A}}))$ with one of the following triplets:
$\left(\widehat{F}_i(\bar{\mathbf{A}}, N, n), W_{a,i}(\bar{\mathbf{A}}), \mathbf{w}_{a,i}(\bar{\mathbf{A}})\right)$ for every $i \in \{1, \ldots, N\}$,
$\left(\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n), W_{a,i}^{(q)}(\bar{\mathbf{A}}), \mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right)$ for every $i \in \{1, \ldots, N\}$,
$\left(\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n), D_w^-(\bar{\mathbf{A}}), \mathbf{d}_w^-(\bar{\mathbf{A}})\right)$, $\left(\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}}, N, n), D_w^{-\,(q)}(\bar{\mathbf{A}}), \mathbf{d}_w^{-\,(q)}(\bar{\mathbf{A}})\right)$, or
$\left(\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n), W_\infty(\bar{\mathbf{A}}), \mathbf{w}_\infty(\bar{\mathbf{A}})\right)$.

We proceed to study the second moment of $X(\bar{\mathbf{A}}, N, n)$. Assuming that each configuration is equally likely, the variance of $X(\bar{\mathbf{A}}, N, n)$ and its limiting behavior as $N \to \infty$ are:

Theorem 1.9 $\operatorname{Var} X(\bar{\mathbf{A}}, N, n) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W(\bar{\mathbf{A}})\right)$. $\operatorname{Var} X(\bar{\mathbf{A}}, N, n) \to 0$ *at rate* N^{-1} *as* $N \to \infty$ *assuming* $\operatorname{Var} W(\bar{\mathbf{A}}) < \infty$.

The variance of the local relative frequency of the attribute directly depends on the population variance of agent weights. If there is large heterogeneity in agents' weights, then the local relative frequency of the attribute strongly varies with configuration, and this gets reflected in the variance of the distribution. This variance is maximal when $f = 0.5$, and it monotonically decreases as f moves away from 0.5. To study the behavior of $\operatorname{Var} X(\bar{\mathbf{A}}, N, n)$ as the population increases in size, introduce replica graphs that both preserve existing agents' relative weights and maintain the amount of weight accorded to a particular indexed node on the graph. By scaling the population upwards in this manner, $\text{Var}\, X\left(\bar{\mathbf{A}}, N, n\right)$ halves as the population size doubles.

In specific settings, we can relate the variance of the weighted local relative frequency of the attribute to network primitives. When $X\left(\bar{\mathbf{A}}, N, n\right) = \widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right)$, $\text{Var}\, \widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right)$ directly depends on $\text{Var}\, D_w^-\left(\bar{\mathbf{A}}\right)$: the capacity for variation in the average local relative frequency of the attribute directly depends on the variance of average weighted in-degrees for the network. Meanwhile, when $X\left(\bar{\mathbf{A}}, N, n\right) = \widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n\right)$, $\text{Var}\, \widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n\right)$ directly depends on the variance of degrees for the network, $\text{Var}\, D\left(\mathbf{A}\right)$, given the network's total number of edges; this relationship holds when graph $\mathcal{G}\left(\mathbf{A}\right)$ is undirected, connected, and aperiodic, and all non-zero elements within every row of $\bar{\mathbf{A}}$ have the same value (see Theorem 1.2).

The closed-form expression for $\text{Var}\, X\left(\bar{\mathbf{A}}, N, n\right)$ in Theorem 1.9 provides us with the necessary mathematics to construct a configuration-induced error bound about the outcome of an aggregated economic system. This error bound quantifies the extent to which there can be variation in the outcome of the economy. The error bound gets constructed about the outcome that results from only considering the aggregate properties of the economy and not the underlying configuration of the attribute. Since agents choose actions based on the local relative frequency of the attribute, we can use $\text{Var}\, X\left(\bar{\mathbf{A}}, N, n\right)$ to compute the extent to which the economy's outcome varies with configuration. The topology of the network determines the size of this error bound.

In the next theorem, we show how to compute the lower and upper bounds on the support of $X\left(\bar{\mathbf{A}}, N, n\right)$. These values represent the lowest and highest possible local relative frequencies of the attribute given the attribute's global relative frequency in the population. They determine the maximal extent to which the local relative frequency of the attribute can deviate from its global relative frequency given the structure of the network. From these values, we are able to bound the distribution of possible outcomes for the economy.

Theorem 1.10 *Construct the ordered multiset* $\{w_s\}_{s=1}^{N}$ *from the elements of* $\mathbf{w}(\bar{\mathbf{A}})$ *so that* $w_s \leq w_{s'}$ *whenever* $s \leq s'$*. The lower and upper bounds on the support of* $X(\bar{\mathbf{A}}, N, n)$ *are respectively:*

$$\min \text{supp } X(\bar{\mathbf{A}}, N, n) = \sum_{s=1}^{n} w_s \text{ and } \max \text{supp } X(\bar{\mathbf{A}}, N, n) = \sum_{s=N-n+1}^{N} w_s.$$

The lower bound on the support of $X(\bar{\mathbf{A}}, N, n)$ is equal to the sum of the n smallest agent weights in the population, while the upper bound is equal to the sum of the n largest agent weights in the population. All possible weighted local relative frequencies of the attribute given f must then fall within these two bounds.

We proceed to study the limiting behavior of $G_{X(\bar{\mathbf{A}},N,n)}(t)$ as $N \to \infty$. To do so, we define the quantity

$$\kappa_N(\epsilon) = \frac{1}{\sum_{i=1}^{N}\left([\mathbf{w}_N(\bar{\mathbf{A}})]_i - \frac{1}{N}\right)^2} \sum_{\substack{j \in \{1,\ldots,N\} \text{ s.t.} \\ \left|[\mathbf{w}_N(\bar{\mathbf{A}})]_j - \frac{1}{N}\right| > \epsilon\sigma_N}} \left([\mathbf{w}_N(\bar{\mathbf{A}})]_j - \frac{1}{N}\right)^2$$

where $\sigma_N = \left(\frac{n}{N}\left(1 - \frac{n}{N}\right)\sum_{i=1}^{N}\left([\mathbf{w}_N(\bar{\mathbf{A}})]_i - \frac{1}{N}\right)^2\right)^{1/2}$. We make the population size, N, explicit for the $N \times 1$ vector $\mathbf{w}_N(\bar{\mathbf{A}})$ of network-derived agent weights because we wish to study the behavior of $G_{X(\bar{\mathbf{A}},N,n)}(t)$ as N increases. We establish the following central limit theorem-type result:

Theorem 1.11 *If* $\lim_{N\to\infty} \kappa_N(\epsilon) = 0$ *for any* $\epsilon > 0$*, then* $\lim_{N\to\infty} G_{\frac{X(\bar{\mathbf{A}},N,n) - \frac{n}{N}}{\sigma_N}}(t) = \Phi(t)$ *for all real* t*, where* $\Phi(\cdot)$ *is the standard normal CDF.*

The requirement that $\lim_{N\to\infty} \kappa_N(\epsilon) = 0$ for any $\epsilon > 0$ is a Lindeberg-type condition.

As the population size increases and the number of agent weights increases, the Lindeberg-type condition requires that there cannot be any subset of agent weights as $N \to \infty$ that strongly deviates from the average agent weight. When this condition holds, we informally have that $\lim_{N\to\infty} G_{X(\bar{\mathbf{A}},N,n)}(t) \approx \Phi\left(\frac{t-\frac{n}{N}}{\sigma_N}\right)$. The distribution of weighted local relative frequencies of the attribute is asymptotically normal with mean $\frac{n}{N}$ and variance σ_N^2, the mean of $X(\bar{\mathbf{A}}, N, n)$ and the variance of $X(\bar{\mathbf{A}}, N, n)$. As the population size increases, provided that the set of agent weights is well-behaved, the population variance of agent weights tends to zero, so $\operatorname{Var} X(\bar{\mathbf{A}}, N, n)$ also tends to zero. As $N \to \infty$, the number of nodes on the network indeed grows as well. The only constraint on the underlying network's growth is that the Lindeberg-type condition continues to be satisfied. From this theoretical result, we see that, as $N \to \infty$, the particular configuration of the binary-valued attribute among agents becomes irrelevant. When every configuration is equally likely, the distribution of possible local relative frequencies of the attribute converges to a degenerate distribution positioned at the attribute's global relative frequency. The rate at which this central limit theorem-type result applies determines the extent to which configuration is still relevant for a large-N population.

The next theorem provides insight into the rate at which $X(\bar{\mathbf{A}}, N, n)$ converges to a normal distribution as the population size increases. It places an upper bound on the maximal distance of the distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$ to a normal distribution with the same mean and variance:[18]

---

[18]Here, we take N to be large enough so that $\operatorname{Var} X(\bar{\mathbf{A}}, N, n) = \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W(\bar{\mathbf{A}})\right) = \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\sum_{i=1}^{N}\left(\left[\mathbf{w}_N(\bar{\mathbf{A}})\right]_i - \frac{1}{N}\right)^2 \approx \frac{n}{N}\left(1-\frac{n}{N}\right)\sum_{i=1}^{N}\left(\left[\mathbf{w}_N(\bar{\mathbf{A}})\right]_i - \frac{1}{N}\right)^2 = \sigma_N^2$.

Theorem 1.12 *For all real* t*,*

$$\left| G_{X(\bar{\mathbf{A}},N,n)}(t) - \Phi\left(\frac{t - \frac{n}{N}}{\sigma_N}\right)\right| \leq \frac{C}{\sqrt{\frac{n}{N}\left(1-\frac{n}{N}\right)}} \frac{\sum_{i=1}^{N}\left|\left[\mathbf{w}_N\left(\bar{\mathbf{A}}\right)\right]_i - \frac{1}{N}\right|^3}{\left(\sum_{i=1}^{N}\left(\left[\mathbf{w}_N\left(\bar{\mathbf{A}}\right)\right]_i - \frac{1}{N}\right)^2\right)^{3/2}},$$

where C *is an absolute constant.*

The upper bound depends on f and the normalized third absolute moment for the distribution of agent weights. It is a Berry-Esseen-type inequality that specifies the rate at which convergence to the normal distribution takes place by bounding the maximal error of approximation.

Beyond the statistical features of $X\left(\bar{\mathbf{A}}, N, n\right)$ provided thus far, we are also interested in the CDF of the distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$. The next result shows, via asymptotic expansion, how we can draw the CDF of our distribution for any feasible population size, network structure, and prevalence of the attribute in the population. Let's begin by defining the function $J\left(\bar{\mathbf{A}}, N, n, t\right)$:

$$\begin{aligned} J\left(\bar{\mathbf{A}}, N, n, t\right) &= \Phi(t) - H_2(t)\,\phi(t)\,C_1\sum_{i=1}^{N}\widehat{w}_i^3 \\ &\quad - H_3(t)\,\phi(t)\left[C_2\left(\sum_{i=1}^{N}\widehat{w}_i^4 - \frac{3}{N}\right) - \frac{1}{4N}\right] - H_5(t)\,\phi(t)\,C_3\left(\sum_{i=1}^{N}\widehat{w}_i^3\right)^2, \end{aligned}$$

where $\widehat{w}_i = \frac{[\mathbf{w}(\bar{\mathbf{A}})]i - EW(\bar{\mathbf{A}})}{\sqrt{N\,\mathrm{Var}\,W(\bar{\mathbf{A}})}}$, $C_1 = \frac{1-\frac{2n}{N}}{6\left(\frac{n}{N}\left(1-\frac{n}{N}\right)\right)^{1/2}}$, $C_2 = \frac{1-6\left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right)}{24\left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right)}$, $C_3 = \frac{\left(1-\frac{2n}{N}\right)^2}{72\left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right)}$, $\phi(t) = \Phi'(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$, and $H_i(t)\,\phi(t) = (-1)^i \frac{d^i}{dt^i}\phi(t)$. Quantity $\widehat{w}_i$ is constructed from the set of agent weights. We can then approximate $G_{\frac{X(\bar{\mathbf{A}},N,n) - EX(\bar{\mathbf{A}},N,n)}{(\mathrm{Var}\,X(\bar{\mathbf{A}},N,n))^{1/2}}}(t)$ by the function $J\left(\bar{\mathbf{A}}, N, n, t\right)$:

Theorem 1.13 *Provided that condition (c) holds,*

$$\left| G_{\frac{X(\bar{\mathbf{A}},N,n) - EX(\bar{\mathbf{A}},N,n)}{(\mathrm{Var}\,X(\bar{\mathbf{A}},N,n))^{1/2}}}(t) - J\left(\bar{\mathbf{A}}, N, n, t\right)\right| < C_4 \times \sum_{i=1}^{N}\left|\widehat{w}_i\right|^5$$

for all t, where C_4 is only a function of $\frac{n}{N}$.

Condition (c) (Robinson (1978)) *Given $C' > 0$, there exist $\epsilon > 0$, $C > 0$, and $\delta > 0$ not depending on N such that, for any fixed t, the number of indices j, for which $\left|\widehat{w}_j \widehat{x} - t - 2r\pi\right| > \epsilon$, for all $\widehat{x} \in \left(C' \left[\max_i |\widehat{w}_i|\right]^{-1}, C \left[\sum_{i=1}^N |\widehat{w}_i|^5\right]^{-1}\right)$ and all $r = 0, \pm 1, \pm 2, \ldots$, is greater than δN, for all N.*

The asymptotic expansion $J(\bar{\mathbf{A}}, N, n, t)$ in Theorem 1.13 is to order $1/N$. Condition (c) requires that the multiset $\{\widehat{w}_i\}_{i=1}^N$ not be clustered around too few values; it therefore also requires that the multiset of agent weights $\{[\mathbf{w}(\bar{\mathbf{A}})]_i\}_{i=1}^N$ not be clustered around too few values. This asymptotic expansion is a general result that enables us to very strongly approximate the distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$, of weighted local relative frequencies of the attribute provided that condition (c) holds:

$$G_{X(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{t - EX(\bar{\mathbf{A}}, N, n)}{(\operatorname{Var} X(\bar{\mathbf{A}}, N, n))^{1/2}}\right).$$

The function $J(\bar{\mathbf{A}}, N, n, t)$ is essentially a collection of terms; the first term is the normal distribution, and the other terms represent deviations away from the normal distribution provided that they are non-zero. Note that $\sum_{i=1}^N \widehat{w}_i^3 = N^{-1/2} \operatorname{Skew} W(\bar{\mathbf{A}})$ and $\sum_{i=1}^N \widehat{w}_i^4 - \frac{3}{N} = N^{-1} \times (\text{Excess Kurtosis}\, W(\bar{\mathbf{A}}))$. Accordingly, we can re-write the function $J(\bar{\mathbf{A}}, N, n, t)$ in terms of the higher-order moments of $W(\bar{\mathbf{A}})$:

$$\begin{aligned} J(\bar{\mathbf{A}}, N, n, t) = \Phi(t) - H_2(t)\,\phi(t)\,C_1 N^{-1/2} \operatorname{Skew} W(\bar{\mathbf{A}}) \\ - H_3(t)\,\phi(t)\left[C_2\left(N^{-1}\,\text{Excess Kurtosis}\, W(\bar{\mathbf{A}})\right) - \frac{1}{4N}\right] \\ - H_5(t)\,\phi(t)\,C_3 N^{-1}\left(\operatorname{Skew} W(\bar{\mathbf{A}})\right)^2. \end{aligned}$$

We can recover the central limit theorem-type result from Theorem 1.11 by noting that if the skewness and kurtosis of $W(\bar{\mathbf{A}})$ are finite, then $\lim_{N\to\infty} J(\bar{\mathbf{A}}, N, n, t) = \Phi(t)$. The extent to which the higher-order moments of the distribution of agent weights are non-zero determines the extent to which $G_{X(\bar{\mathbf{A}},N,n)}(t)$ deviates from a normal distribution.

We now demonstrate how skewness of $W(\bar{\mathbf{A}})$ directly generates skewness of $X(\bar{\mathbf{A}}, N, n)$. Take the derivative of $J(\bar{\mathbf{A}}, N, n, t)$ with respect to t to find an approximating probability density function to $g_{\frac{X(\bar{\mathbf{A}},N,n)-EX(\bar{\mathbf{A}},N,n)}{(\text{Var}\,X(\bar{\mathbf{A}},N,n))^{1/2}}}(t)$:[19]

$$J'(\bar{\mathbf{A}}, N, n, t) \equiv \frac{\partial J(\bar{\mathbf{A}}, N, n, t)}{\partial t} = \phi(t) + H_3(t)\,\phi(t)\,C_1 \sum_{i=1}^{N} \widehat{w}_i^3$$
$$+ H_4(t)\,\phi(t)\left[C_2\left(\sum_{i=1}^{N} \widehat{w}_i^4 - \frac{3}{N}\right) - \frac{1}{4N}\right] + H_6(t)\,\phi(t)\,C_3\left(\sum_{i=1}^{N} \widehat{w}_i^3\right)^2.$$

The second and fourth terms in the expansion depend on the skewness of $W(\bar{\mathbf{A}})$. If we expand the second term, we find that it is an odd function. Provided that $f < 0.5$ and Skew $W(\bar{\mathbf{A}}) > 0$, the second term reallocates mass away from the normal density function $\phi(t)$ to generate positive skewness. If $f > 0.5$ and Skew $W(\bar{\mathbf{A}}) > 0$, the second term reallocates mass away from the normal density function $\phi(t)$ to generate negative skewness. The more heavily skewed the set of agent weights, the more heavily skewed $X(\bar{\mathbf{A}}, N, n)$. Even though the fourth term depends on the skewness of $W(\bar{\mathbf{A}})$, it is an even function, so the reallocation of mass away from the normal distribution has no effect on skewness.

Skewness of $X(\bar{\mathbf{A}}, N, n)$ matters, particularly when the distribution is unimodal, because it determines the extent to which the median of the distribution deviates from the mean of the distribution. Then, the probability that the local

[19] The distance, $\left| g_{\frac{X(\bar{\mathbf{A}},N,n)-EX(\bar{\mathbf{A}},N,n)}{(\text{Var}\,X(\bar{\mathbf{A}},N,n))^{1/2}}}(t) - J'(\bar{\mathbf{A}}, N, n, t) \right|$, can also be bounded from above. See line (14) of Robinson (1978), for example.

relative frequency of the attribute is greater than f does not equal the probability that the local relative frequency of the attribute is less than f. As a result, given a random configuration, the network topology might be such that it favors relatively higher or relatively lower local relative frequencies of the attribute. Depending on the particular setting and the particular real-world interpretation of the binary-valued attribute, this deviation of the mean from the median can be important. In Appendix A.4, we explore in more detail how the higher-order features of the distribution of agent weights shape the higher-order features of $G_{X(\bar{\mathbf{A}},N,n)}(t)$. The relationship between kurtosis of $W(\bar{\mathbf{A}})$ and kurtosis of $X(\bar{\mathbf{A}},N,n)$ is a bit more complicated.

We see that the distributional features of $W(\bar{\mathbf{A}})$ shape the distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$; that relationship becomes explicit when we examine the function $J(\bar{\mathbf{A}},N,n,t)$. Now, the vector of agent weights is itself network-derived, so ultimately, it is the topological features of agents' interaction network that shape the distributional features of $G_{X(\bar{\mathbf{A}},N,n)}(t)$. When $X(\bar{\mathbf{A}},N,n) = \widehat{F}_{avg}(\bar{\mathbf{A}},N,n)$, for example, the shape of the CDF $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ depends on the statistical moments for the distribution of average weighted in-degrees. The distribution of average weighted in-degrees therefore determines the shape of the distribution of possible outcomes for the economy. Meanwhile, when $X(\bar{\mathbf{A}},N,n) = \widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)$, the shape of $G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(t)$ depends on the statistical moments for the distribution of degrees, provided that $\mathcal{G}(\mathbf{A})$ is undirected, connected, and aperiodic, and all non-zero elements within every row of $\bar{\mathbf{A}}$ have the same value (see Theorem 1.2). For this case, the shape of the degree distribution determines the shape of the distribution of outcomes for the economy.

We can now revisit Example 1.3 from Section 1.3 and show how to compute by hand the probability that Trump's expected vote share exceeds 0.5, or equivalently, the probability that the average local unemployment rate exceeds 0.10:

$$\begin{aligned}\Pr\left[\widehat{F}_{avg}\left(\bar{\mathbf{A}},N,n\right)>0.10\right]&=1-G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}\left(0.10\right)\\&\approx 1-J\left(\bar{\mathbf{A}},N,n,\frac{0.10-E\widehat{F}_{avg}\left(\bar{\mathbf{A}},N,n\right)}{\left(\mathrm{Var}\,\widehat{F}_{avg}\left(\bar{\mathbf{A}},N,n\right)\right)^{1/2}}\right)\\&=0.0707,\text{where } \widehat{w}_i=\frac{\left[\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)\right]_i-ED_w^-\left(\bar{\mathbf{A}}\right)}{\left(N\,\mathrm{Var}\,D_w^-\left(\bar{\mathbf{A}}\right)\right)^{1/2}}.\end{aligned}$$

We have a closed-form approximation for $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}\left(t\right)$. The relevant vector of agent weights here is the vector of average weighted in-degrees, $\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)$.

Theorem 1.13 is quite flexible; it allows us to draw the CDF $G_{X(\bar{\mathbf{A}},N,n)}\left(t\right)$ for every feasible population size, network topology, and prevalence of the attribute. It will also later enable us to provide a closed-form expression for the distribution of outcomes for the economy. However, if agents' weights are clustered over too few values, then condition (c) does not hold and Theorem 1.13 no longer applies. We therefore proceed to provide a theoretical result that allows us to construct $g_{X(\bar{\mathbf{A}},N,n)}\left(t\right)$ in closed form in certain settings when we are unable to apply the findings of Theorem 1.13.

We consider an environment in which k_ω agents have the same non-zero weight ω and $N-k_\omega$ agents have zero weight. If we define the set $\mathcal{I}$,

$$\mathcal{I}=\left\{\max\left\{0,n-\left(N-k_\omega\right)\right\},\max\left\{0,n-\left(N-k_\omega\right)\right\}+1,\ldots,\min\left\{n,k_\omega\right\}\right\},$$

we then have the following result:

Theorem 1.14 *For all $i\in\mathcal{I}$,*

$$g_{X(\bar{\mathbf{A}},N,n)}\left(i\omega\right)=\frac{\binom{k_\omega}{i}\binom{N-k_\omega}{n-i}}{\binom{N}{n}}.$$

For all $i\notin\mathcal{I}$, $g_{X(\bar{\mathbf{A}},N,n)}\left(i\omega\right)=0$.

In such a setting, the probability mass function is almost everywhere zero. There are only finitely many values for which it is non-zero. Those values are integer multiples of ω, where the set of allowable integers i is restricted to those in set $\mathcal{I}$. The probability that the local relative frequency of the attribute equals $i\omega$ is then equal to the fraction of all possible configurations such that there are i individuals with weight ω and the attribute's unit value and $n-i$ individuals with weight zero and the attribute's unit value. When agent j equally weights each of his linkages on the network, the resulting vector of agent weights, $\mathbf{w}_{a,j}(\bar{\mathbf{A}})$, takes this exact form and Theorem 1.14 applies for constructing $g_{\widehat{F}_{a,j}(\bar{\mathbf{A}},N,n)}(t)$.

Appendix A.4 provides some additional results concerning the properties of $G_{X(\bar{\mathbf{A}},N,n)}(t)$. Appendix A.5 characterizes the statistical features of the multivariate random variable $\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}},N,n)$, whose realizations are $\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}},\mathbf{b},N,n)$, the population vector of weighted local relative frequencies of the attribute. Appendix A.6 identifies those vectors of agent weights and corresponding network topologies for which $\operatorname{Var} X(\bar{\mathbf{A}},N,n)$ is maximal. Appendix A.7 characterizes those vectors of agent weights and those matrices, $\bar{\mathbf{A}}$, that generate identical distributions $G_{X(\bar{\mathbf{A}},N,n)}(t)$ and therefore identical distributions of outcomes for an economy. Lastly, Appendix A.8 conducts sensitivity analysis, studying how a perturbation to agents' interaction network affects $G_{X(\bar{\mathbf{A}},N,n)}(t)$ via its effects on the relevant network-derived vector of agent weights.

1.6.2 Non-Degeneracy of the Distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$ for Very Large N

We now examine how $G_{X(\bar{\mathbf{A}},N,n)}(t)$ and its mathematical analogue $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ can remain approximately non-degenerate even as the population size grows very large.

In Section 1.3, we studied the distribution, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, of possible average local unemployment rates for a population of 137.5 million voters, given an October 2016 U-6 unemployment rate of 9.6 percent and an interaction network characterized by the composite graph. We observed strong configuration dependence of the average local unemployment rate for very large N, with a standard deviation of 0.266 percentage points. We identified two different ways for understanding such strong configuration dependence: (1) high variance of in-degrees for the composite graph, and (2) heavy-tailedness in the distribution of agent weights, that is, heavy-tailedness in the distribution of average weighted in-degrees, $G_{D_{\bar{w}}^{-}(\bar{\mathbf{A}})}(t)$, for the composite graph. We now show how properties (1) and (2) of the composite graph generate a strongly non-degenerate distribution, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, even for large N.

We first show how high variance of in-degrees for the composite graph generates an approximately non-degenerate distribution, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, of possible average local unemployment rates. To demonstrate this relationship, we must slightly modify the structure of the composite graph. We require voters to assign an equal weight to each of their observations, so that non-zero elements for each row of $\bar{\mathbf{A}}$ take the same value; this assumption happens to be one that we were already making when carrying out analysis in Section 1.3. We also require every agent to have the same out-degree, k, setting $k = \frac{1}{N}\mathbf{1}^T\mathbf{d}^-(\mathbf{A})$.[20] This latter assumption is reasonable given that the standard deviation for the distribution of out-degrees is just 18.4, while the standard deviation for the distribution of in-degrees is 8,633.3. The next theorem makes explicit the relationship between the properties of the in-degree distribution for the slightly modified composite graph and the properties of $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$:

[20] We allow k to be non-integral.

Theorem 1.15 *Let $\mathbf{d}^+(\mathbf{A}) = k\mathbf{1}$ and let all non-zero matrix elements in each row of $\bar{\mathbf{A}}$ take the same value. Then, $\mathbf{d}_w^-(\bar{\mathbf{A}}) = \frac{1}{Nk}\mathbf{d}^-(\mathbf{A})$, so*

$$\operatorname{Var}\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}N\left(\frac{1}{Nk}\right)^2 \operatorname{Var} D^-(\mathbf{A}).$$

Construct the ordered multiset $\{v_s\}_{s=1}^N$ from the elements of $\mathbf{d}^-(\mathbf{A})$ so that $v_s \leq v_{s'}$ whenever $s \leq s'$. The lower and upper bounds on the support of $\widehat{F}_{avg}(\bar{\mathbf{A}})$ are respectively

$$\min\operatorname{supp}\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n) = \frac{1}{Nk}\sum_{s=1}^{n} v_s \ \ \textit{and} \ \ \max\operatorname{supp}\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n) = \frac{1}{Nk}\sum_{s=N-n+1}^{N} v_s.$$

Every quantity of interest in Theorem 1.15 directly depends on the composite graph's in-degree distribution. For a fixed number of edges, the greater the variance of the graph's in-degrees, the greater the variance of $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$. We compute $\operatorname{Var}\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ and the bounds on the support for $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ using the expressions from Theorem 1.15. We find that the standard deviation of $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ is 0.00307, the minimal bound on the support of $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ is 0.0528, and the maximal bound on the support of $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ is 0.366. These values are quite similar to the ones that we computed exactly before, with a standard deviation of 0.00266, a minimal bound of 0.0553, and a maximal bound of 0.333. The approximate relation between $\mathbf{d}_w^-(\bar{\mathbf{A}})$ and $\mathbf{d}^-(\mathbf{A})$ for the composite graph explains why the shape of $G_{D_w^-(\bar{\mathbf{A}})}(t)$ in Figure 1.8 is very similar to the shape of $G_{D^-(\mathbf{A})}(t)$ in Figure 1.7.

With $\widehat{w}_i = \frac{\left[\mathbf{d}_w^-(\bar{\mathbf{A}})\right]_i - ED_w^-(\bar{\mathbf{A}})}{\left(N\operatorname{Var}D_w^-(\bar{\mathbf{A}})\right)^{1/2}}$, we can approximate $G_{\frac{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n) - E\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}{\left(\operatorname{Var}\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)\right)^{1/2}}}(t)$ by the function $J(\bar{\mathbf{A}}, N, n, t)$:

$$\begin{aligned} J(\bar{\mathbf{A}}, N, n, t) = \Phi(t) &- H_2(t)\,\phi(t)\,C_1\left(N^{-1/2}\operatorname{Skew}D^-(\mathbf{A})\right) \\ &- H_3(t)\,\phi(t)\left[C_2\left(N^{-1}\text{Excess Kurtosis}\,D^-(\mathbf{A})\right) - \frac{1}{4N}\right] \\ &- H_5(t)\,\phi(t)\,C_3\left(N^{-1}\left(\operatorname{Skew}D^-(\mathbf{A})\right)^2\right), \end{aligned}$$

with quantities C_1, C_2, $C_{3,}$, $\phi(t)$, and $H_i(t)\phi(t)$ defined in Theorem 1.13. The distribution of in-degrees shapes the CDF $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, and it therefore shapes the distribution of outcomes for the economy.

We now remove the previous two restrictions on the structure of the composite graph. We show, in general, how heavy-tailedness in a distribution of agent weights generates a distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$, that is non-degenerate even for very large N. We introduce random variable $\widetilde{W}_i(\bar{\mathbf{A}})$ with non-negative support. $\widetilde{W}_i(\bar{\mathbf{A}})$ denotes the effective representation of agent i in the population. This mass equals 1 on average for each agent, hence $E\widetilde{W}_i(\bar{\mathbf{A}}) = 1$. Here, we assume that random variables $\widetilde{W}_i(\bar{\mathbf{A}})$ for $i \in \{1, \ldots, N\}$ are independent and identically distributed.[21] The CDF from which the $\widetilde{W}_i(\bar{\mathbf{A}})$'s are drawn is $G_{\widetilde{W}(\bar{\mathbf{A}})}(t) = G_{W(\bar{\mathbf{A}})\times N}(t)$, the distribution of network-derived agent weights scaled by N. $X(\bar{\mathbf{A}}, N, n)$ is constructed by drawing n values $\widetilde{W}_i(\bar{\mathbf{A}})$ from $G_{\widetilde{W}(\bar{\mathbf{A}})}(t)$ and designating those agents as the ones with the attribute's unit value: $X(\bar{\mathbf{A}}, N, n) = \frac{\widetilde{W}_1(\bar{\mathbf{A}})+\cdots+\widetilde{W}_n(\bar{\mathbf{A}})}{N}$. We then have the following result:

Theorem 1.16 *Suppose that* $\text{Var}\,\widetilde{W}_i(\bar{\mathbf{A}})$ *is finite. Then as* $n, N \to \infty$ *holding* $f = \frac{n}{N}$ *fixed,*

$$n^{1/2}\left(\frac{1}{f}X(\bar{\mathbf{A}}, N, n) - E\widetilde{W}_i(\bar{\mathbf{A}})\right) \xrightarrow[d]{} \mathcal{N}\left(0, \text{Var}\,\widetilde{W}_i(\bar{\mathbf{A}})\right)$$

where $E\widetilde{W}_i(\bar{\mathbf{A}}) = 1$. *Next suppose that* $\Pr\left[\widetilde{W}_i(\bar{\mathbf{A}}) > t\right] \sim L(t)\,t^{-\xi}$, *where* $L(t)$ *is a slowly-varying function and* $\xi \in (1,2)$. *Then as* $n, N \to \infty$ *holding* $f = \frac{n}{N}$ *fixed,*

$$n^{1-1/\xi}\left(\frac{1}{f}X(\bar{\mathbf{A}}, N, n) - E\widetilde{W}_i(\bar{\mathbf{A}})\right) \xrightarrow[d]{} \widetilde{S}(\xi, \beta, \widetilde{\gamma}, 0; 1),$$

where $\widetilde{S}(\xi, \beta, \widetilde{\gamma}, 0; 1)$ *is a stable distribution and* $E\widetilde{W}_i(\bar{\mathbf{A}}) = 1$.

---

[21] In reality, the random variables $\widetilde{W}_i(\bar{\mathbf{A}})$ are not independent, as they are constrained to sum to N: $\widetilde{W}_1(\bar{\mathbf{A}}) + \cdots + \widetilde{W}_N(\bar{\mathbf{A}}) = N$.

When $\xi \in (1,2)$, agent weights are power-law distributed with a finite mean and infinite variance.[22] Empirically estimating the power-law exponent, we find that $\xi \approx 1.69 \in (1,2)$. This exponent emerges from the right tail of the counter-cumulative distribution function of agents' effective representations in the population, where $\widetilde{W}(\bar{\mathbf{A}}) = W(\bar{\mathbf{A}}) \times N$ and an agent's weight is equal to that agent's average weighted in-degree for the composite graph. With $n^{1-1/\xi} < n^{1/2}$ for all $\xi \in (1,2)$, the distribution for $X(\bar{\mathbf{A}}, N, n)$ collapses faster to a degenerate distribution as $N \to \infty$ when agent weights have a finite variance than when they are drawn from a power-law distribution with infinite variance. The rate at which the law of large numbers applies is relatively slower in this latter case, so the variance of $X(\bar{\mathbf{A}}, N, n)$ becomes non-negligible even at extremely large population sizes, N. Accordingly, in the setting of Section 1.3, distribution $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ is approximately degenerate for N = 137.5 million when the underlying interaction network is the base graph, but it is approximately non-degenerate for N = 137.5 million when the underlying interaction network is the composite graph.

We see that economies with all possible population sizes can be configuration dependent. Even in large-N settings, the aggregate properties of the system are not sufficient to determine how the economy will evolve. We need to account for the configuration dependence of the system so that we can construct the entire distribution of possible outcomes.

---

[22]The literature tends to use α instead of ξ, but the former variable has already been assigned a different interpretation in this chapter.

1.7 Features of the Precursor Distribution When Configurations are Not Equally Likely

We continue to characterize the precursor distribution of possible local relative frequencies of the attribute given the economy's aggregate feature, $f = \frac{n}{N}$. However, we relax the assumption that every configuration $\mathbf{b}(N, n) \in \mathcal{B}(N, n)$ of the attribute is equally likely. Before, when each configuration of the attribute was equally likely, every agent i had the same probability of $b_i = 1$. Now, every agent i in the system has an arbitrary probability that $b_i = 1$, so configurations can occur with any relative likelihood. In this relaxed setting, we solve for the first two moments of the resulting probability distribution of local relative frequencies of the attribute.

We let each agent i have a vector of characteristics, γ_i, that can impact $\phi_i = \Pr[B_i = 1 | \gamma_i]$, the conditional probability that agent i has the binary-valued attribute's unit value. B_i is a random variable whose realization is agent i's binary-valued attribute: 0 or 1. We partition agent indices into Θ categories according to their conditional probabilities, so that agents i, j are in category θ if $\phi_i = \phi_j = \rho_\theta$. We define the odds ratio for agents in category θ relative to category k as follows: $\widehat{\psi}_\theta = \frac{\frac{\rho_\theta}{1-\rho_\theta}}{\frac{\rho_k}{1-\rho_k}}$, with $\widehat{\psi}_k \equiv 1$. Then:

Theorem 1.17 *The first two moments of* $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ *are:*

$$EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = \sum_{i=1}^N [\mathbf{w}(\bar{\mathbf{A}})]_i [\boldsymbol{\mu}]_i$$
$$\text{and } \operatorname{Var} X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = [\mathbf{w}(\bar{\mathbf{A}})]^T \boldsymbol{\Sigma} [\mathbf{w}(\bar{\mathbf{A}})].$$

To compute the $N \times 1$ *vector* $\boldsymbol{\mu}$ *and the* $N \times N$ *matrix* $\boldsymbol{\Sigma}$*, define the* $\Theta \times 1$ *vector* $\widehat{\boldsymbol{\mu}}$ *across the* Θ *categories; set* $[\boldsymbol{\mu}]_i = [\widehat{\boldsymbol{\mu}}]_\theta$ *for each agent* i *from category* θ*. Also introduce the* $N \times 1$ *vector* $\boldsymbol{\zeta}$*, setting* $[\boldsymbol{\zeta}]_i = [\widehat{\boldsymbol{\mu}}]_\theta (1 - [\widehat{\boldsymbol{\mu}}]_\theta)$ *for each agent* i *from category* θ*. Define the* $\Theta \times \Theta$

matrix $\widehat{\mathbf{\Sigma}}$ *with element* $\left[\widehat{\mathbf{\Sigma}}\right]_{\theta k}$ *equal to the conditional covariance* $\operatorname{Cov}\left(B_i, B_j\right)$ *between agent* i *in category* θ *and agent* j *in category* k *and element* $\left[\widehat{\mathbf{\Sigma}}\right]_{\theta\theta}$ *equal to the conditional variance* $\operatorname{Var} B_i$ *for agent* i *in category* θ. $\boldsymbol{\mu}$ *and* $\mathbf{\Sigma}$ *can be approximated by solving the following system of equations:*

$$\sum_{\theta=1}^{\Theta} \sum_{\substack{i\in\{1,\ldots,N\} \\ s.t.\ \phi_i=\rho_\theta}} [\widehat{\boldsymbol{\mu}}]_\theta = n$$

$$\widehat{\psi}_\theta = \frac{[\widehat{\boldsymbol{\mu}}]_\theta \left(1-[\widehat{\boldsymbol{\mu}}]_k\right) - \left[\widehat{\mathbf{\Sigma}}\right]_{\theta k}}{\left(1-[\widehat{\boldsymbol{\mu}}]_\theta\right)[\widehat{\boldsymbol{\mu}}]_k - \left[\widehat{\mathbf{\Sigma}}\right]_{\theta k}}, \quad \forall\theta\in\{1,\ldots,\Theta\}\setminus\{k\},\ \textit{and}$$

$$\mathbf{\Sigma} = \frac{N}{N-1}\left(\operatorname{diag}\boldsymbol{\zeta} - \frac{\boldsymbol{\zeta}\boldsymbol{\zeta}^T}{\mathbf{1}^T\boldsymbol{\zeta}}\right).$$

The $N \times 1$ random vector $\mathbf{B}$, whose i^{th} element is random variable B_i, is distributed according to Fisher's multivariate non-central hypergeometric distribution. $\boldsymbol{\mu} = E\mathbf{B}$ is the $N \times 1$ conditional mean vector for $\mathbf{B}$ and $\mathbf{\Sigma}$ is the $N \times N$ conditional covariance matrix for $\mathbf{B}$. $\mathbf{B}$, $\boldsymbol{\mu}$, and $\mathbf{\Sigma}$ are quantities that correspond to the population of N agents. We can also introduce parallel hatted quantities that correspond to the Θ distinct categories. If we define a $\Theta \times 1$ random vector $\widehat{\mathbf{B}}$, whose θ^{th} element is random variable B_i from category θ, then $\widehat{\boldsymbol{\mu}} = E\widehat{\mathbf{B}}$ is the corresponding $\Theta \times 1$ conditional mean vector for $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{\Sigma}}$ is the corresponding $\Theta \times \Theta$ conditional covariance matrix for $\widehat{\mathbf{B}}$; diagonal element $\left[\widehat{\mathbf{\Sigma}}\right]_{\theta\theta}$ is the conditional variance $\operatorname{Var} B_i$ for agent i in category θ. We also introduce the $\Theta \times 1$ vector $\widehat{\mathbf{\Sigma}}^{Cov}$; the θ^{th} element $\left[\widehat{\mathbf{\Sigma}}^{Cov}\right]_\theta$ of this vector is the conditional covariance $\operatorname{Cov}\left(B_i, B_j\right)$ for agent i and agent j, $i \neq j$, both in category θ.

When the population of agents can be partitioned into Θ categories, the total number of variables and the total number of equations in the system both equal $2\Theta + \binom{\Theta}{2} + \sum_{\theta=1}^{\Theta} \mathbb{1}_{s_\theta>1}$, where s_θ is the number of agents in category θ. There are Θ variables $[\widehat{\boldsymbol{\mu}}]_\theta$, the conditional mean EB_i for agent i in category θ. There are Θ

variables $\left[\widehat{\boldsymbol{\Sigma}}\right]_{\theta\theta}$, the conditional variance Var B_i for agent i in category θ. There are $\binom{\Theta}{2}$ variables $\left[\widehat{\boldsymbol{\Sigma}}\right]_{\theta\theta'}$ for $\theta \neq \theta'$, the conditional covariance Cov $\left(B_i, B_j\right)$ for agent i in category θ and agent j in category θ'. Lastly, there are $\sum_{\theta=1}^{\Theta} \mathbb{1}_{s_\theta > 1}$ variables $\left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{\theta}$; when there is more than one agent in category θ, we must also compute the conditional covariance Cov $\left(B_i, B_j\right)$ for agent i and agent j in category θ. We construct $\boldsymbol{\mu}$ from the elements of $\widehat{\boldsymbol{\mu}}$ and we construct $\boldsymbol{\Sigma}$ from the elements of $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\Sigma}}^{Cov}$.

When $\phi_i = \Pr\left[B_i = 1 \middle| \gamma_i\right]$ differs across agents, $EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ no longer must equal the global relative frequency of the attribute, f. The weighted local relative frequency of the attribute can, on average, be either greater or less than f. In the social learning setting, when configurations are not equally likely, it may not be possible for there to be asymptotic convergence to the truth, as the mean of the distribution of $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ need not equal f for all iterations of learning q and all population sizes N, even as $q, N \to \infty$. When configurations are no longer equally likely to occur, we can induce bias in $EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ away from the global relative frequency, f.

The next example works through the null case, in which every configuration is equally likely; it computes the first two moments of $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ when $\Theta = 1$, following Theorem 1.17:

Example 1.7 (First Two Moments of $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$, $\Theta = 1$) *Consider an economic system with N agents. When $\Theta = 1$, so that $\phi_i = \rho_1$ for every agent $i \in \{1, \ldots, N\}$, $EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = EX\left(\bar{\mathbf{A}}, N, n\right)$ and* Var $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) =$ Var $X\left(\bar{\mathbf{A}}, N, n\right)$, *where $X\left(\bar{\mathbf{A}}, N, n\right)$ is the random variable of interest when every configuration* *$\mathbf{b}(N, n) \in \mathcal{B}(N, n)$ is equally likely.*

When $\Theta = 1$ we recover the first two moments $EX\left(\bar{\mathbf{A}}, N, n\right)$ and Var $X\left(\bar{\mathbf{A}}, N, n\right)$

that we studied in Section 1.6 when every configuration was equally likely. The corresponding derivation is in Appendix A.9.

In the more general setting, we can characterize these first two moments for particular cases of $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$, such as $\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ and $\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$; we simply follow Theorem 1.17 and make the necessary substitutions. Appendix A.9 studies three different economies with $\Theta \neq 1$. The first economy features $N = 4$ agents and $\Theta = 2$ categories of agents; we compute the first two moments of $\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$. The second economy features $N = 15$ agents and $\Theta = 3$ categories of agents; we compute the first two moments of $\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$. The third economy revisits the setting of Section 1.3 and Example 1.4, in which there are $N = 137.5$ million voters deciding which 2016 U.S. presidential candidate to elect. Here, we have $\Theta = 2$ categories of agents: (1) those featured by the media, and (2) those not featured by the media. Setting $\rho_1 = 0.50$ and $\rho_2 = 0.096$, Appendix A.9 shows how to compute $E\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = 0.194$ and $\text{Std. Dev.}\,\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = 0.00452$, the first two moments of the distribution of possible average local unemployment rates given that the actual unemployment rate, f, is 0.096. In all settings, the first two moments of the distribution deviate from the ones that would occur when configurations are equally likely.

1.8 The Distribution of Outcomes for the Economy

Thus far, we have developed the mathematics that enables us to characterize the distributional features of $X\left(\bar{\mathbf{A}}, N, n\right)$ and $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$. In this section, we compute the distribution of possible outcomes for the economy. For certain classes of agent actions, we can provide a closed-form expression for the distribution of possible outcomes. We already demonstrated in Section 1.3 how to compute such a distribution of possible outcomes for the economy. In that setting, agents made a voting decision based on their locally formed macroeconomic sentiments; we computed both the probability that the election outcome favored candidate Clinton and the probability that the election outcome favored candidate Trump. In this section, we show how to construct the distribution of possible outcomes for the economy when agents follow other decision-making rules. The classes of agent decision-making rules that we identify and study in this section are certainly not exhaustive.

Before we can study the distribution of possible outcomes for the economy, we must introduce some additional notation. Constructing this distribution requires us to incorporate agents' decision-making behavior, so much of the notation concerns agent actions. We assume that each agent i chooses an action, $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, that depends on his or her local relative frequency of the attribute: $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = h_i(x_i(\bar{\mathbf{A}}, \mathbf{b}, N, n))$. $h_i(\cdot)$ is a function that specifies how agent i responds to his or her weighted local relative frequency of the attribute, $x_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, where $x_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{w}_i(\bar{\mathbf{A}})]^T \mathbf{b}(N, n)$. Since agents may respond to different local relative frequencies of the same attribute, we index by agent the quantity $x_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and the vector of agent weights $\mathbf{w}_i(\bar{\mathbf{A}})$.

We are interested in the individual action $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and how it varies with configuration given f, and we are also interested in the population's aggregate action $a_{agg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \sum_{i=1}^{N} a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and how that quantity varies with configuration given f. We define random variables $A_i(\bar{\mathbf{A}}, N, n)$ and $A_{agg}(\bar{\mathbf{A}}, N, n)$ with respective configuration-specific realizations $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and $a_{agg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$. In this section, we study the distributions of $A_i(\bar{\mathbf{A}}, N, n)$ and $A_{agg}(\bar{\mathbf{A}}, N, n)$. When each configuration of an attribute is equally likely to occur, random variable $A_i(\bar{\mathbf{A}}, N, n)$

has CDF: $G_{A_i(\bar{\mathbf{A}},N,n)}(t) = \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{a_i(\bar{\mathbf{A}},\mathbf{b},N,n)\le t}$, and random variable $A_{agg}\left(\bar{\mathbf{A}}, N, n\right)$ has CDF: $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t) = \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{a_{agg}(\bar{\mathbf{A}},\mathbf{b},N,n)\le t}$.

1.8.1 The Distribution of Outcomes Given the Action of Agent i

We compute the distribution of possible outcomes for the economy when the relevant object is the action of agent i. Therefore, we study agent i's action, $a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, and we seek to characterize the distribution $G_{A_i(\bar{\mathbf{A}},N,n)}\left(t\right)$. We focus on the case in which agents' weights are such that they satisfy condition (c) of Theorem 1.13. We can write an expression for the distribution $G_{A_i(\bar{\mathbf{A}},N,n)}\left(t\right)$ provided that agent i's action, $a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = h_i\left(x_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right)$, is invertible over the domain $x_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) \in [0, 1]$.[23] Our expression for $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ is then:

$$G_{A_i(\bar{\mathbf{A}},N,n)}(t) = G_{X_i(\bar{\mathbf{A}},N,n)}\left(h_i^{-1}\left(t\right)\right) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{h_i^{-1}\left(t\right) - EX_i\left(\bar{\mathbf{A}}, N, n\right)}{\left(\operatorname{Var} X_i\left(\bar{\mathbf{A}}, N, n\right)\right)^{1/2}}\right),$$

where $EX_i\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}$, $\operatorname{Var} X_i\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W_i\left(\bar{\mathbf{A}}\right)\right)$, and $\widehat{w}_j = \frac{\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_j - EW_{a,i}(\bar{\mathbf{A}})}{\left(N \operatorname{Var} W_{a,i}(\bar{\mathbf{A}})\right)^{1/2}}$. When agent i's action takes an affine form:

$$a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = h_i\left(x_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right) = \alpha_i x_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) + \beta_i,$$

[23] To see why there exists a closed-form expression for $G_{A_i(\bar{\mathbf{A}},N,n)}\left(t\right)$ when agent i's action, $a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = h_i\left(x_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right)$ is invertible, follow the derivation below:

$$\begin{aligned} G_{A_i(\bar{\mathbf{A}},N,n)}(t) &= \Pr\left[A_i\left(\bar{\mathbf{A}}, N, n\right) \le t\right] \\ &= \Pr\left[h_i\left(X_i\left(\bar{\mathbf{A}}, N, n\right)\right) \le t\right], \text{ with } \operatorname{supp} X_i\left(\bar{\mathbf{A}}, N, n\right) \in [0, 1] \\ &= \Pr\left[X_i\left(\bar{\mathbf{A}}, N, n\right) \le h_i^{-1}(t)\right] \\ &= G_{X_i(\bar{\mathbf{A}},N,n)}\left(h_i^{-1}\left(t\right)\right) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{h_i^{-1}\left(t\right) - EX_i\left(\bar{\mathbf{A}}, N, n\right)}{\left(\operatorname{Var} X_i\left(\bar{\mathbf{A}}, N, n\right)\right)^{1/2}}\right). \end{aligned}$$

action $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is invertible with respect to $x_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, so we can express $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ in closed form: $G_{A_i(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{\frac{t-\beta_i}{\alpha_i} - EX_i(\bar{\mathbf{A}},N,n)}{(\text{Var } X_i(\bar{\mathbf{A}},N,n))^{1/2}}\right) = J\left(\bar{\mathbf{A}}, N, n, \frac{t-EA_i(\bar{\mathbf{A}},N,n)}{(\text{Var } A_i(\bar{\mathbf{A}},N,n))^{1/2}}\right)$. By computing the mean action for agent i, $EA_i(\bar{\mathbf{A}}, N, n) = \alpha_i \frac{n}{N} + \beta_i$, and we see that the distribution of possible outcomes for the economy, $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$, is centered about the outcome in which configuration is irrelevant and only the aggregate feature of the economy, $f = \frac{n}{N}$, matters. Appendix A.10 features two examples that solve for $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ and its distributional characteristics in closed form. For the first example, agent i's action is an affine transformation of the weighted local relative frequency of the attribute, $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \alpha_i \widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \beta_i$, and for the second example, agent i's action nonlinearly depends on the weighted local relative frequency of the attribute, $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \alpha_i \log \widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \beta_i$. In both settings, we can write a closed-form expression for $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$, the distribution of possible outcomes for the economy, for every feasible population size, network topology, and prevalence of the binary-valued attribute's unit value in the population.

1.8.2 The Distribution of Outcomes Given the Aggregate Action

We now compute the distribution of possible outcomes for the economy when the relevant object is the aggregate action. We therefore study the aggregate action, $a_{agg}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, with an interest in characterizing the distribution $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$. We assume that agents' weights are such that they satisfy condition (c) of Theorem 1.13. To characterize $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$ in closed form, we focus on particular classes of agent actions.

Our main class of agent actions is the one in which individual agents' actions take an affine form: $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \alpha_i \widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \beta_i, \ \forall i \in \{1, \ldots, N\}$.

We can then write a closed-form expression for $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$ for every feasible population size, underlying network topology, and global relative frequency of the attribute. Observe that the mean aggregate action, $EA_{agg}(\bar{\mathbf{A}},N,n)$, and therefore the mean outcome for the economy, is the one that occurs when the system is invariant to configuration and only the aggregate feature of the system matters. The extent to which the aggregate action can deviate away from this mean aggregate action determines how dependent the system is on configuration. When $\alpha_i = \alpha$ for all agents, $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{t-\alpha NE\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)-\mathbf{1}^T\boldsymbol{\beta}}{\alpha N\left(\text{Var}\,\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)\right)^{1/2}}\right)$, with $J(\cdot)$ defined in Theorem 1.13 and $\widehat{w}_i = \frac{\left[\mathbf{d}_{\overline{w}}^{-}(\bar{\mathbf{A}})\right]_i - ED_{\overline{w}}^{-}(\bar{\mathbf{A}})}{\left(N\,\text{Var}\,D_{\overline{w}}^{-}(\bar{\mathbf{A}})\right)^{1/2}}$. When agents no longer have a common coefficient α, $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{t-\left(\mathbf{1}^T\boldsymbol{\alpha}\right)E\widehat{F}_{avg}\left(\widehat{\mathbf{A}},N,n\right)-\mathbf{1}^T\boldsymbol{\beta}}{\left(\mathbf{1}^T\boldsymbol{\alpha}\right)\left(\text{Var}\,\widehat{F}_{avg}\left(\widehat{\mathbf{A}},N,n\right)\right)^{1/2}}\right)$, with $J(\cdot)$ defined in Theorem 1.13, $\widehat{w}_i = \frac{\left[\widehat{\mathbf{d}}_{\overline{w}}^{-}\left(\widehat{\mathbf{A}}\right)\right]_i - E\widehat{D}_{\overline{w}}^{-}\left(\widehat{\mathbf{A}}\right)}{\left(N\,\text{Var}\,\widehat{D}_{\overline{w}}^{-}\left(\widehat{\mathbf{A}}\right)\right)^{1/2}}$, $\left[\widehat{\mathbf{A}}\right]_{ij} = \alpha_i\left[\bar{\mathbf{A}}\right]_{ij}$, and $\widehat{\mathbf{d}}_{\overline{w}}^{-}\left(\widehat{\mathbf{A}}\right) = \left(\frac{1}{\alpha_1+\cdots+\alpha_N}\right)\widehat{\mathbf{A}}^T\mathbf{1}$. Appendix A.10 solves for $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$ and its distributional features for these two settings, when agents have a common coefficient α and when agents no longer have a common coefficient α.

When agents' actions instead depend on the consensus local relative frequency of the attribute, there exists an even larger class of decision rules for which we can write a closed-form expression for $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$. If agent actions follow a threshold rule:

$$a_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right) = \begin{cases} \beta_i & \text{if } \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right) \geq \alpha \\ 0 & \text{if } \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}},\mathbf{b},N,n\right) < \alpha \end{cases}, \text{ then}$$

$$A_{agg}\left(\bar{\mathbf{A}},N,n\right) \approx \begin{cases} \mathbf{1}^T\boldsymbol{\beta} & \text{with probability } 1 - J\left(\bar{\mathbf{A}}, N, n, \frac{\alpha - E\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}{\left(\text{Var}\,\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)\right)^{1/2}}\right) \\ 0 & \text{with probability } J\left(\bar{\mathbf{A}}, N, n, \frac{\alpha - E\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}{\left(\text{Var}\,\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)\right)^{1/2}}\right) \end{cases},$$

with $J(\cdot)$ defined in Theorem 1.13 and $\widehat{w}_i = \frac{\left[\mathbf{w}_{\infty}(\bar{\mathbf{A}})\right]_i - EW_{\infty}(\bar{\mathbf{A}})}{\left(N\,\text{Var}\,W_{\infty}(\bar{\mathbf{A}})\right)^{1/2}}$.

Even if we cannot express $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$ in closed form, we can still solve for its lower-order features. If agent i's action, $\forall i \in \{1,\ldots,N\}$, follows a different threshold rule:

$$a_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right) = \begin{cases} \beta_i & \text{if } \widehat{f}_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right) \geq \alpha \\ 0 & \text{if } \widehat{f}_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right) < \alpha \end{cases},$$

we can solve for $EA_{agg}\left(\bar{\mathbf{A}},N,n\right)$ in closed form:

$$EA_{agg}\left(\bar{\mathbf{A}},N,n\right) \approx \sum_{i=1}^{N}\beta_i\left[1 - J\left(\bar{\mathbf{A}},N,n,\frac{\alpha - E\widehat{F}_i\left(\bar{\mathbf{A}},N,n\right)}{\left(\operatorname{Var}\widehat{F}_i\left(\bar{\mathbf{A}},N,n\right)\right)^{1/2}}\right)\right],$$

with $J(\cdot)$ defined in Theorem 1.13 and $\widehat{w}_j = \frac{\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_j - EW_{a,i}(\bar{\mathbf{A}})}{\left(N\operatorname{Var}W_{a,i}(\bar{\mathbf{A}})\right)^{1/2}}$. Appendix A.10 studies in greater detail these two cases in which agent actions follow a threshold rule.

Given agents' decision-making behavior, we are able to both construct the distribution of possible outcomes for the economy and study the distribution's statistical features. We can assess the non-degeneracy of the distribution of outcomes and therefore determine the extent to which the economy is dependent on configuration. When the outcome of the economy strongly varies with configuration, the economy's aggregate feature, f, is no longer sufficient for determining how the economy will evolve, and agents' local interactions and local environments matter for the behavior of the overall economy. For various classes of agent decision-making rules, we can write closed-form expressions for $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ and $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$ for any feasible population size, network topology, and aggregate feature, and we can study how the statistical features of the distributions $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ and $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$ emerge from the topology of agents' interaction network.

1.9 Conclusion

In this chapter, we develop a set of mathematical tools that allows us to map network structures to probability distributions. Given agents' decision-making behavior, we are able to map the specific topology of agents' interaction network to a specific probability distribution of possible outcomes for the economy. We first map the structure of agents' network, $\mathcal{G}(\bar{\mathbf{A}})$, to a precursor distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$, of possible local relative frequencies of the binary-valued attribute. Then, given agents' decision-making behavior and the precursor distribution, $G_{X(\bar{\mathbf{A}},N,n)}(t)$, we construct the distribution of possible outcomes for the economy.

Our characterization of $G_{X(\bar{\mathbf{A}},N,n)}(t)$ is complete. For all population sizes, feasible network structures, and possible global prevalences of the attribute in the population, we can solve for the distributional features of $G_{X(\bar{\mathbf{A}},N,n)}(t)$ in closed form, and we can provide a closed-form expression for the actual shape of this probability distribution. Our mathematical tools show how network primitives and other features of the underlying network directly generate probability distributions with certain statistical features. Meanwhile, for particular classes of agent actions, we can write a closed-form expression for the distribution of possible outcomes for the economy as well; for a larger set of agent actions, we can solve for the lower-order features of this distribution in closed form. When this distribution is non-degenerate, the particular configuration of the attribute among agents matters for how the economy evolves.

The tools and content developed in this work have several implications. First, these mathematical tools enable closed-form analysis of complex economic systems. Such tools allow us to map complex agent interactions into a simple probability distribution characterizing how the system will probabilistically evolve. Second, these mathematical tools allow us to quantify how dependent an aggregate economic system is on the underlying configuration of an attribute among its agents. Aggregate treatments of economic systems put forth a single outcome for the economy, but if the economic system is configuration dependent, there is actually an entire non-degenerate distribution of possible outcomes for the economy consistent with the aggregate properties of the system. The tools of this work allow us to construct in closed form an error bound about the original benchmark outcome of the aggregate economy. Third, the theoretical framework of this work and the accompanying mathematical tools help us to understand locally formed macroeconomic sentiment. This work offers a microfoundation for animal spirits, showing how agents' interaction structure enables the existence of swings in aggregate sentiment for fixed economic fundamentals that persist even for large-N economies. Chapter 2 of this dissertation utilizes and extends the mathematical tools and theoretical framework developed in this work for an entirely different economic application. Hopefully the theoretical tools and methodology developed in the present chapter can be broadly used to provide insights in diverse settings.

Chapter 2

The Distribution of Multipliers in a Networked Economy

2.1 Introduction

There are many economic settings in which agents are networked and the actions that agents take are interdependent. The complexities of such network-based agent interactions can make it difficult to ascertain in advance the effects of a planned policy on agents' aggregate behavior. Given that it is difficult to ascertain the aggregate action in advance, it is also difficult to predict the policy-specific economic multiplier; this latter quantity measures the change in the aggregate action arising from implementation of the planned policy. The present chapter shows how to compute the effects of a planned policy on the aggregate action and how to determine the policy's economic multiplier. For each policy, given agents' decision-making behavior and the topology of agents' interaction network, there are entire probability distributions of possible aggregate actions and economic multipliers. This chapter develops and applies a set of theoretical tools so that we can explicitly map the planned policy to the corresponding probability distributions of possible aggregate actions and economic multipliers.

In this work, we focus on a population of N networked agents, each of whom chooses an action, and an outside actor. The outside actor is interested in the population's aggregate action. In particular, the outside actor would like to adjust the aggregate action and therefore chooses a policy with that intention. We can imagine that this actor would like to increase the aggregate action above its no-intervention level. There are many different real-world settings that parallel the theoretical setting of the present work, with its N networked agents and outside actor interested in increasing the aggregate action. The outside actor might be a government interested in jump-starting its economy during a recession; the government would like to provide stimulus to a set of firms organized on a production network with the intention of increasing aggregate output. The outside actor might alternatively be a nonprofit organization, such as a cancer research foundation; the foundation would like to increase the amount of innovative activity in cancer research. To achieve this goal, it allocates funds to research groups who are organized on both formal and informal R&D networks; the linkages of these networks capture both collaboration and competition among research groups.

There are many different classes of policies that the outside actor can implement. In this work, we focus on one type of policy. Here, when a policy gets enacted, the outside actor transmits a positive shock of ϵ magnitude to $n \leq N$ agents. Such a shock can be a positive wealth shock, in which the outside actor provides $\epsilon > 0$ units of additional wealth to n agents. In general, though, the exact interpretation of the shock depends on the environment. There are two different ways that the outside actor can finance its policy. The outside actor can either gather funds from the other networked agents in the population, or it can receive funds from agents who are outside of the system. We refer to the former case as a setting with internally provided transfers, and we refer to the latter case as a setting with externally provided stimulus. The internal transfers might be implemented via taxation, while the external stimulus might originate from issuance of debt or donation. In the setting with transfers, the financing agents receive a negative shock, so the net adjustment that the policy initially induces for all agents in the population is zero.

When issuing a policy, the outside actor chooses both n and the method of financing. Therefore, given a particular policy, we can introduce a binary-valued attribute that identifies which agents have received a positive shock; we assign the attribute's unit value to the n agents targeted by the policy, while we assign the attribute's zero value to the other $N - n$ agents. This assignment of positive shocks to n agents represents a particular configuration of positive shocks; specifically, the configuration identifies the indices for the subset of n agents who have the attribute's unit value. Given that a policy targets n agents, there are $\binom{N}{n}$, or combinatorially many, total possible configurations of positive shocks. For each configuration, a different group of agents receives the positive shock. Agents' actions are interdependent, so the aggregate action and economic multiplier can vary depending on which group of agents actually receives the positive shock. Accordingly, holding both n and the method of financing fixed, we can construct an entire distribution of possible aggregate actions and an entire distribution of possible economic multipliers.

When the outside actor is planning to issue a policy targeting n agents, the actor is interested in the full range of possible outcomes. As a result, both the distribution of possible aggregate actions and the distribution of possible economic multipliers are the relevant theoretical objects of interest. We can alternatively imagine that the outside actor knows the set of agents that it would like to target, but it does not know which nodes on the network these agents occupy. In such circumstances, the distribution of possible aggregate actions and the distribution of possible economic multipliers constructed from all feasible configurations are also the appropriate objects of interest. We can separately imagine that the outside actor selects agents at random for receipt of a positive shock; the outside actor would then like to know the distribution of possible aggregate actions and the distribution of possible economic multipliers as well. Across all scenarios, the outside actor has a sense of the topology for the underlying agent interaction network. If the outside actor does not quite know the underlying network's topology, it can carry out sensitivity analysis, perturbing different features of the underlying network and examining how these perturbations shift the policy-induced distributions of possible aggregate actions and economic multipliers.

In this work, we develop a large body of theoretical results. These results explicitly show how the outside actor's policy generates the probability distributions of possible aggregate actions and economic multipliers. We are able to map the outside actor's policy to two simple probability distributions even though agents' interactions are complex. Given agents' decision-making behavior, we can characterize these probability distributions for every feasible policy, that is, for every feasible number of agents being targeted by the policy, n, and for both methods of financing. The shapes of these probability distributions fundamentally depend on the population size of networked agents, N, the number of agents being targeted by the policy, n, and the topology of agents' interaction network; our theoretical results explicitly show how each of these quantities impacts the statistical features for our distributions.

We characterize in closed form the main statistical features as well as the

CDFs for both the distribution of aggregate actions and the distribution of economic multipliers. We solve for the first two moments of these distributions. We find that, in settings with transfers, the mean aggregate action is always equal to its no-intervention level and the mean economic multiplier is always equal to zero. On average, the outside actor's policy has no effect. In settings with stimulus, however, the mean aggregate action can deviate from its no-intervention level, and the mean economic multiplier can deviate from zero. We also present closed-form expressions for the second moments of these distributions. These second moments determine how much risk is entailed in enacting a particular policy. We find that the second moments for these distributions depend on the variance of average weighted in-degrees for a network that is a mathematical transformation of the original agent interaction structure. Beyond these first two moments, we provide closed-form expressions for the lower and upper bounds on the supports of the distributions of aggregate actions and economic multipliers. Given a particular policy, we are able to determine both the worst and the best possible outcomes, and we can show how these values depend on the topology of the underlying network. We develop a theoretical result that essentially allows us to draw the CDFs for the distributions of aggregate actions and economic multipliers. From this result, we can see how the topology of the mathematically transformed network generates properties of skewness and/or heavy-tailedness in the distributions of aggregate actions and economic multipliers. If these distributions are heavy-tailed, extreme values for the aggregate action and economic multiplier can be more likely than the outside actor might have otherwise thought. Meanwhile, in the setting with transfers, skewness in the distribution of economic multipliers might mean that the probability of a negative multiplier is greater than the probability of a positive multiplier, which can make implementation of the policy less attractive. We can characterize the distributions of aggregate actions and economic multipliers in the limit as $N \to \infty$ as well. All of these theoretical results equip the policy-making actor with the tools that it might need to evaluate the effects and pitfalls of implementing particular policies in networked environments.

We develop and present all of these theoretical results for a general networked environment. We then proceed by studying three specific environments with network-based interaction: (1) networked environments with strategic complements and strategic substitutes, (2) networked environments with coordination and anti-coordination, and (3) networked environments with production. All of the results from the general environment apply to these three broad classes of environments. When we develop the general set of results, we introduce two free parameters into the theory: a matrix and a scalar quantity. For each of the three networked environments, specific expressions for these two parameters naturally emerge. The general theoretical results therefore nest the specific theoretical results for each of the three networked environments, and the mathematics ends up being the same. For each specific environment, we develop additional theoretical results. We identify the topologies of those networks for which the distributions of aggregate actions and economic multipliers are exactly degenerate; then, both the aggregate action and the economic multiplier are invariant to configuration. When possible, we identify the network structures that deliver the highest feasible economic multiplier and the lowest feasible economic multiplier. We moreover rank networks so that the distributions of aggregate actions and economic multipliers for relatively higher-ranked networks first-order stochastically dominate the distributions of aggregate actions and economic multipliers for relatively lower-ranked networks. Consequently, the higher-ranked the network, the more effective the policy.

Quite importantly, for both the general networked environment and the three specific networked environments, we show how to compute in closed form the probability that the aggregate action ends up being below its no-intervention level and the economic multiplier ends up being negative. The outside actor chooses to implement a policy with the intention of increasing the aggregate action, but depending on the topology of agents' interaction network, there can be a non-negligible probability that the policy ends up being harmful. Indeed, in settings with transfers, provided that agents' weights are not all equal, there is *always* a positive probability of a negative multiplier for every level n; agents' weights here are equal to the average weighted in-degrees for the mathematically transformed network. Whenever agents' interaction network is non-trivial, negative multipliers emerge naturally, especially in settings with transfers but also in settings with stimulus. By being able to compute the probability of a negative multiplier in closed form, the outside actor can better assess the risks inherent in enacting a particular policy.

2.1.1 Relation to the Literature

This work interfaces with four different areas of the literature: research on (1) networks, (2) economic complexity, (3) economic multipliers, and (4) fiscal stimulus. Throughout the present work, agent interaction is organized on networks. In particular, for environments featuring strategic complements and strategic substitutes, and environments featuring coordination and anti-coordination, agents choose actions by playing games on networks. The recent literature on network games includes work by Ballester, Calvó-Armengol, and Zenou (2006), Galeotti et al. (2010), Jackson and Yariv (2011), Jackson and Zenou (2015), and Jackson, Rogers, and Zenou (2017). The present work uses the structure of network games and the flexibility that they offer to study aggregate actions in a variety of environments and characterize how aggregate actions adjust once the behavior of some subset of agents is perturbed. The present work also studies environments with production networks. Recent literature on production networks includes Acemoglu et al. (2012), Chaney (2014), Acemoglu, Akcigit, and Kerr (2016), Barrot and Sauvagnat (2016), Boehm, Flaaen, and Pandalai-Nayar (2017), and Oberfield (2018). Several of these papers study the transmission of shocks across production networks. In the present paper, shocks are demand-originating; rather than the value of the shock be stochastic, as in other work, shocks in the present work are fixed in magnitude and they target a subset of firms and/or sectors. Randomness here arises in the particular configuration of firms and/or sectors actually receiving a positive shock given that the implemented policy exactly targets n firms and/or sectors. Our object of interest is the *probability distribution* of possible levels of aggregate output that result once a fixed number of firms and/or sectors are targeted via a demand channel.

The theoretical environment of this work is complex. The present work therefore engages strongly with past research on economic complexity, some of which includes Topa (2001), Brock and Durlauf (2001a), and Brock and Durlauf (2001b). These past papers all develop interactions-based models, with interactions either being local or global social interactions. The present work focuses on local interactions; it examines agent decision-making when interactions are defined locally by an underlying network structure. Complexity in the present work emerges when we look at the behavior of the aggregate action and the corresponding economic multiplier after a policy targets n agents. Network-based interactions cause the aggregate action to adjust more or less than the aggregate action would absent any network-based interaction. The networked system moreover exhibits a degree of nonlinearity. Given that a group of n agents has received an initial positive shock, as we incrementally increase the additional agents receiving that positive shock, the sequence of changes in the aggregate action is, in general, very much nonlinear. The present work also departs from past work in the area of economic complexity by studying probability distributions. Rather than identifying a unique equilibrium, this work characterizes probability distributions of possible equilibria, focusing on aggregate actions in the economy and corresponding economic multipliers. We can study how the structure of agents' interaction network shapes the distribution of aggregate actions and the distribution of economic multipliers. As in Chapter 1 of this dissertation, the present chapter condenses the complexities of agent-based interactions into a probability distribution. Chapter 1 operates in a general setting; it maps networks and agent actions to a probability distribution of possible outcomes for the economy, given the global prevalence of some binary-valued attribute within the population. The present work extends and applies the theoretical results from Chapter 1, mapping networks and agents' decision-making behavior to a distribution of possible aggregate actions and economic multipliers; the binary-valued attribute here denotes an agent's receipt of a positive transfer or positive stimulus.

The two main objects studied in the present work are the aggregate action and the corresponding economic multiplier. We define the economic multiplier as the increase in the aggregate action that results when n agents each receive an additional unit of a positive transfer or positive stimulus. The economic multiplier therefore varies with configuration, and given n, we can construct the entire distribution of possible economic multipliers. Recent research concerning economic multipliers focuses on social multipliers and network multipliers. Glaeser, Sacerdote, and Scheinkman (2003) and Calvó-Armengol and Zenou (2004) identify social multipliers. They study the disconnect between aggregate-level behavior and individual-level behavior, attributing that disconnect to social interactions;

these two works therefore study the social multiplier effects that exist on individual decisions. Baqaee (2013), Carvalho (2014), and Acemoglu, Akcigit, and Kerr (2016) meanwhile identify network multipliers. Carvalho (2014) and Acemoglu, Akcigit, and Kerr (2016) study how the network structure of the economy amplifies sector-specific volatility. Baqaee (2013) introduces an employment multiplier that studies adjustments to equilibrium employment following sector-specific shocks; the employment multiplier depends on the underlying structure of the network. The present work enriches the existing literature on multipliers by studying and characterizing entire distributions of multipliers whose values depend on agents' interaction structure and the particular class of agent actions. This work additionally introduces closed-form expressions that allow us to compute the probability that an economic multiplier is negative.

Now, the present work studies an outside entity whose policy provides positive transfers or positive stimulus to n agents in the system. If this outside entity is indeed a government, then the types of policies that we are examining are fiscal policies, and the positive shocks to the n agents in the system represent fiscal stimulus. Research in the area of fiscal stimulus includes: Christiano, Eichenbaum, and Rebelo (2011), Woodford (2011), Eggertsson (2011), Ilzetzki, Mendoza, and Végh (2013), Nakamura and Steinsson (2014), Farhi and Werning (2016), Chodorow-Reich (2018), and Hagedorn, Manovskii, and Mitman (2018). The Great Recession and the handicapping of monetary policy during the zero-lower-bound environment of the time sparked renewed interest in fiscal multipliers. On the theory side, research has focused on mapping macroeconomic models with particular underlying assumptions and features to corresponding expressions for the government expenditure multiplier. Such models include variants of neoclassical models, New Keynesian models, open-economy models, closed-economy models, representative agent models, and heterogeneous agent models. Depending on the particular environment and how government spending is financed, the magnitude of the fiscal multiplier differs. On the empirical side, recent work has sought to estimate the value of the government spending multiplier in different environments; Nakamura and Steinsson (2014) estimate an open-economy multiplier, and Chodorow-Reich (2018) bounds the national government expenditure multiplier from estimates of cross-sectional fiscal spending multipliers. These multipliers need not always be positive. Ilzetzki, Mendoza, and Végh (2013) empirically estimate negative fiscal multipliers in countries with high public debt ratios. Hagedorn, Manovskii, and Mitman (2018) show theoretically that tax-financed fiscal stimulus generates multipliers whose magnitudes are smaller than those of fiscal multipliers financed externally via issuance of debt. When the outside actor in our work is a government, the economic multipliers that we study end up being fiscal multipliers. For a given level of fiscal stimulus, which is either financed internally by tax or transfer or financed externally, we can compute the entire non-degenerate distribution of possible fiscal multipliers. The particular configuration of fiscal stimulus among economic agents fundamentally matters. Through this work, we also introduce a new channel by which fiscal multipliers can be negative: network-based interactions among agents in the economy. We can illustrate how negative fiscal multipliers arise from natural patterns of agent interaction, and we can quantify the probability that a particular level of fiscal stimulus decreases aggregate output or the aggregate action below its no-intervention level.

2.1.2 Outline of Chapter

Section 2.2 begins by introducing notation and definitions. It then proceeds to develop a high-level unifying theoretical framework for the study of policy-induced distributions of aggregate actions and economic multipliers in a general networked setting. The theoretical framework of Section 2.2 nests the specific networked environments of Sections 2.3, 2.4, and 2.5. Section 2.3 focuses on distributions of aggregate actions and economic multipliers in networked environments with strategic complements and strategic substitutes. Section 2.4 characterizes the policy-induced distributions of aggregate actions and economic multipliers in networked environments with coordination and anti-coordination. Section 2.5 studies the distributions of aggregate output and economic multipliers in networked environments with production. Section 2.6 concludes.

2.2 Theoretical Framework

2.2.1 Notation and Definitions

The cardinality of a set $\mathcal{X}$ is $|\mathcal{X}|$. A *multiset* is an object similar to a set, but it allows for multiple instances of each of its elements. Vector $\mathbf{x}$ is a column vector by default. The i^{th} element of vector $\mathbf{x}$ is x_i or $[\mathbf{x}]_i$. The ij^{th} element of matrix $\mathbf{X}$ is $[\mathbf{X}]_{ij}$, the i^{th} row of $\mathbf{X}$ is $[\mathbf{X}]_{i*}$ and the j^{th} column of $\mathbf{X}$ is $[\mathbf{X}]_{*j}$. $\mathbf{x}' \geq \mathbf{x}$ for vectors $\mathbf{x}, \mathbf{x}'$ if $[\mathbf{x}']_i \geq [\mathbf{x}]_i$ element-wise; meanwhile, $\mathbf{x}' > \mathbf{x}$ if $[\mathbf{x}']_i \geq [\mathbf{x}]_i$ element-wise with at least one integer i for which $[\mathbf{x}']_i > [\mathbf{x}]_i$. $\mathbf{X}' \geq \mathbf{X}$ for matrices $\mathbf{X}, \mathbf{X}'$ if $[\mathbf{X}']_{ij} \geq [\mathbf{X}]_{ij}$ for all pairs (i, j); meanwhile, $\mathbf{X}' > \mathbf{X}$ if $[\mathbf{X}']_{ij} \geq [\mathbf{X}]_{ij}$ element-wise with at least one pair (i, j) for which $[\mathbf{X}']_{ij} > [\mathbf{X}]_{ij}$. The identity matrix is $\mathbf{I}$ and the column vector whose elements all equal 1 is $\mathbf{1}$. Depending on the particular context, $\mathbf{0}$ is either a column vector whose elements all equal 0, or a matrix whose elements all equal 0. $\mathbb{Z}_+$ is the set of all non-negative integers.

Matrix $\mathbf{X}$ is *row-stochastic* if $\mathbf{X}\mathbf{1} = \mathbf{1}$ and all matrix elements of $\mathbf{X}$ are non-negative. Matrix $\mathbf{X}$ is *column-stochastic* if $\mathbf{X}^T\mathbf{1} = \mathbf{1}$ and all matrix elements of $\mathbf{X}$ are non-negative. Matrix $\mathbf{X}$ is *doubly stochastic* if it is both row-stochastic and column-stochastic. The Hadamard product of matrices $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{X} \circ \mathbf{Y}$, is their element-wise multiplication: $[\mathbf{X} \circ \mathbf{Y}]_{ij} = [\mathbf{X}]_{ij} [\mathbf{Y}]_{ij}$. Non-negative matrix $\mathbf{X}$ is *primitive* if there exists an integer $q \geq 1$ such that $[\mathbf{X}^q]_{ij} > 0$ for all matrix elements in $\mathbf{X}^q$. Matrix $\mathbf{X}$ is *semi-convergent* if the limit $\lim_{q\to\infty} \mathbf{X}^q$ exists. $(\lambda, \mathbf{w})$ is a left eigenpair of matrix $\mathbf{X}$ if $\mathbf{w}^T\mathbf{X} = \lambda\mathbf{w}^T$; $(\lambda, \mathbf{w})$ is the *dominant* left eigenpair of $\mathbf{X}$ when the magnitude $|\lambda|$ weakly exceeds that of all other eigenvalues of $\mathbf{X}$. For permutation matrix $\mathbf{P}$, $\mathbf{PX}$ permutes the rows of $\mathbf{X}$ and $\mathbf{XP}$ permutes the columns of $\mathbf{X}$. The *spectral radius* $r(\mathbf{X})$ of a matrix $\mathbf{X}$ is the largest absolute value among the eigenvalues of $\mathbf{X}$. Real random variable $X' \succeq X$ *in the usual stochastic order* if $\Pr[X' > t] \geq \Pr[X > t]$ for all $t \in (-\infty, \infty)$; random variable X' then *first-order stochastically dominates* random variable X.

Graph[1] $\mathcal{G}$ is an ordered pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consisting of a set of vertices (nodes) $\mathcal{V}$ and a set of edges $\mathcal{E}$. $(i, j, e_{i,j}) \in \mathcal{E}$ is an edge between nodes i and j with weight $e_{i,j}$. If the graph is directed, the edge is oriented from node i to node j; otherwise, the edge is not oriented. If we define a weighted adjacency matrix $\mathbf{X}$ whose ij^{th} element $[\mathbf{X}]_{ij} = e_{i,j}$ denotes the edge weight between nodes i and j, then $\mathcal{G}(\mathbf{X}) = (\mathcal{V}(\mathbf{X}), \mathcal{E}(\mathbf{X}))$ is the corresponding weighted graph.

2.2.2 Theoretical Preliminaries

In this work, we study three different environments featuring N networked agents: (1) networked environments with strategic complements and strategic substitutes, (2) networked environments with coordination and anti-coordination, and (3) net-

[1]We will use the terms *graph* and *network* interchangeably.

worked environments with production. In each environment, an outside actor prescribes a policy that exogenously delivers a positive shock of a fixed magnitude to the actions of $n \leq N$ agents. The policy targets n agents, but it does not specify the identities of those agents. The policy is either financed internally by the other agents in the population or it is financed externally by agents who are outside of the system. We therefore have two classes of policies; we refer to the former class of policies as transfers and the latter class of policies as stimulus. Given a prescribed policy, we are interested in the resulting distributions of possible aggregate actions and corresponding economic multipliers. We would like to understand how these distributions and their statistical properties depend on the implemented policy and the topology of the network.

We find that the mathematics is the same across all three environments. Therefore, in the present section, we write out general expressions for the aggregate action and the corresponding economic multiplier for each of our two classes of policies. These general expressions for the aggregate action and the economic multiplier depend on a matrix $\mathbf{Z}$ and parameter γ_1. In later sections, we identify the specific expressions for $\mathbf{Z}$ and γ_1 for each of the three networked environments. We substitute those quantities into the general expressions for the aggregate action and the economic multiplier provided in Section 2 to obtain environment-specific formulae.

In this section, we solve for all distributional features of the aggregate action and the economic multiplier using the general expressions. Later on, we can solve for the distributional features of the aggregate action and the economic multiplier in specific environments by plugging in the relevant values of $\mathbf{Z}$ and γ_1. These results that characterize the distributional features of the aggregate action and the corresponding economic multiplier build on a theoretical framework and a core set of mathematics introduced and developed in Chapter 1 of this dissertation. We adapt the theoretical framework and the core set of mathematical results for the present paper. The present subsection develops this core set of mathematical results as a collection of lemmata. All of the general results characterizing the distributions of aggregate actions and economic multipliers, which build on the set of lemmata, are then presented in the next subsection.

We have a population of N agents organized on a network $\mathcal{G}(\mathbf{Z}) = (\mathcal{V}(\mathbf{Z}), \mathcal{E}(\mathbf{Z}))$ with weighted adjacency matrix $\mathbf{Z}$. There are no constraints on the matrix elements of $\mathbf{Z}$, other than that each element is a real number. Now, every agent i in the population has a binary-valued attribute, b_i. This binary-valued attribute identifies which agents in the population are recipients of positive funds; depending on the policy, these funds are either financed by transfer or stimulus. Specifically, attribute $b_i = 1$ if agent i is the recipient of positive funds, and otherwise $b_i = 0$. There are $n \leq N$ agents receiving positive funds and therefore $n \leq N$ agents with the attribute's unit value. Given n, there is a particular configuration, or arrangement, of this binary-valued attribute among agents in the population. A *configuration* is defined as follows:

Definition 2.1 *A* configuration $\mathbf{b} \equiv \mathbf{b}(N, n)$ *of a binary-valued attribute in a population of N agents is an allocation of this attribute so that $b_i \in \{0, 1\}$ for all agents $i \in \{1, \ldots, N\}$ and $\mathbf{b}^T\mathbf{1} = n$.*

A configuration $\mathbf{b} \equiv \mathbf{b}(N, n)$ is an allocation of the attribute's unit value to exactly n agents in a total population of N agents. Vector $\mathbf{b}$ stacks each agent's binary-valued attribute and identifies the indices of those agents that have the attribute's unit value. Two configurations $\mathbf{b}, \mathbf{b}'$ are distinct when $\mathbf{b} \neq \mathbf{b}'$ because the subsets of agents with the attribute's unit value differ across these two configurations. Given a population of size N and n agents with the attribute's unit value, there are many different possible configurations of the attribute. The set of all possible configurations is $\mathcal{B}(N, n)$, and the cardinality of this set is combinatorial: $|\mathcal{B}(N, n)| = \binom{N}{n}$.

For each agent positioned on the network, we compute the local relative frequency of the attribute's unit value. Since the attribute's unit value denotes the receipt of positive funds, we are therefore essentially computing the local prevalence of this positive shock for every agent in his or her network neighborhood. We compute this quantity for each agent by taking a weighted sum of the values of the binary attribute for the agent's out-neighbors. The weights that we use in this sum are the edge weights that link the agent to these out-neighbors. We accordingly define $\widehat{\mathbf{f}}(\mathbf{Z}, \mathbf{b}, N, n) = \mathbf{Z}\mathbf{b}(N, n)$ as the $N \times 1$ population vector of local relative frequencies of the attribute; the local relative frequency of the attribute for each agent depends on the topology of the network, $\mathcal{G}(\mathbf{Z})$, and it depends on which subset of n agents actually has the attribute's unit value, $\mathbf{b}(N, n)$. The local relative frequency of the attribute for agent i, $\widehat{f}_i(\mathbf{Z}, \mathbf{b}, N, n) = [\mathbf{Z}]_{i*}\mathbf{b}(N, n)$, can take values outside of the $[0, 1]$ interval because $[\mathbf{Z}]_{i*}$ is not constrained to have all non-negative elements and $[\mathbf{Z}]_{i*}\mathbf{1}$ is not constrained to sum to 1. In Chapter 1 of this dissertation, matrix $\mathbf{Z} \equiv \bar{\mathbf{A}}$ is row-stochastic, so all of its elements are non-negative and the elements in each row sum to 1. As a result, for all agents $i \in \{1, \ldots, N\}$, $[\bar{\mathbf{A}}]_{i*}\mathbf{b}(N, n) \in [0, 1]$, and we can exactly interpret $[\bar{\mathbf{A}}]_{i*}\mathbf{b}(N, n)$ as the local relative frequency of the attribute for agent i. To be consistent with the nomenclature of Chapter 1, we also refer to $\widehat{f}_i(\mathbf{Z}, \mathbf{b}, N, n) = [\mathbf{Z}]_{i*}\mathbf{b}(N, n)$ as the local relative frequency of the attribute for agent i; if we want to be more precise, we can think of this quantity as a *scaled* local relative frequency of the attribute.

We are interested in the population-averaged local relative frequency of the attribute's unit value, $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$. We are essentially computing the average local prevalence of the positive shock given n. We calculate this quantity as follows:

$$\widehat{f}_{avg}\left(\mathbf{Z},\mathbf{b},N,n\right)=\frac{1}{N}\mathbf{1}^{T}\widehat{\mathbf{f}}\left(\mathbf{Z},\mathbf{b},N,n\right)=\frac{1}{N}\mathbf{1}^{T}\mathbf{Z}\mathbf{b}\left(N,n\right)=\left[\mathbf{d}_{w}^{-}\left(\mathbf{Z}\right)\right]^{T}\mathbf{b}\left(N,n\right),$$

where $\mathbf{d}_{w}^{-}\left(\mathbf{Z}\right)=\frac{1}{N}\mathbf{Z}^{T}\mathbf{1}$ is the vector of average weighted in-degrees for graph $\mathcal{G}\left(\mathbf{Z}\right)$. To determine the average local relative frequency of the attribute for a particular configuration, we derive the vector of agent weights, $\mathbf{d}_{w}^{-}\left(\mathbf{Z}\right)$, from the underlying graph $\mathcal{G}\left(\mathbf{Z}\right)$. We then multiply $\mathbf{d}_{w}^{-}\left(\mathbf{Z}\right)$ by the configuration vector $\mathbf{b}\left(N,n\right)$, and we obtain the configuration-specific average local relative frequency of the attribute's unit value. The higher an agent's weight, the more that agent contributes to the average local relative frequency of the attribute provided that he or she has the attribute's unit value. The sum of agents' weights is k: $\mathbf{1}^{T}\mathbf{d}_{w}^{-}\left(\mathbf{Z}\right)=k$. Without restrictions on $\mathbf{Z}$, the average local relative frequency of the attribute is not constrained to the interval $[0,1]$; in Chapter 1, $\widehat{f}_{avg}\left(\mathbf{Z},\mathbf{b},N,n\right)\in[0,1]$ because $\mathbf{Z}\equiv\bar{\mathbf{A}}$ is row-stochastic. To be consistent with the nomenclature of Chapter 1, we refer to $\widehat{f}_{avg}\left(\mathbf{Z},\mathbf{b},N,n\right)$ as the average local relative frequency of the attribute; if we want to be more precise, we can think of this quantity $\widehat{f}_{avg}\left(\mathbf{Z},\mathbf{b},N,n\right)$ as a *scaled* average local relative frequency of the attribute.

Holding n fixed, that is, holding fixed the total number of agents receiving positive funds, we can imagine that $\widehat{f}_{avg}\left(\mathbf{Z},\mathbf{b},N,n\right)$ varies with configuration. Depending on which subset of agents receives the positive shock, we can have variation in its average local prevalence. We would like to determine the distribution of possible average local relative frequencies of the positive shock given N, n, and $\mathbf{Z}$. We therefore introduce random variable $\widehat{F}_{avg}\left(\mathbf{Z},N,n\right)$. This random variable has configuration-specific realizations $\widehat{f}_{avg}\left(\mathbf{Z},\mathbf{b},N,n\right)$. The CDF of $\widehat{F}_{avg}\left(\mathbf{Z},N,n\right)$ is $G_{\widehat{F}_{avg}\left(\mathbf{Z},N,n\right)}\left(t\right)$, and we are interested in the distributional features of $\widehat{F}_{avg}\left(\mathbf{Z},N,n\right)$. We assume that each configuration is equally likely, although this is an assumption that we can relax (see Chapter 1). By characterizing the distributional features of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$, we are able to later compute in closed form the distributional features of the aggregate action and corresponding economic multiplier across all three networked environments.

We now present a set of lemmata that characterizes the distributional features of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$. We begin with the first moment of the distribution:

Lemma 2.1 $E\widehat{F}_{avg}(\mathbf{Z}, N, n) = \frac{kn}{N}$.

The average local relative frequency of the attribute can vary with configuration, but across all possible configurations, this quantity is equal to $\frac{kn}{N}$ on average. The value of constant k depends on the topology of network $\mathcal{G}(\mathbf{Z})$.

We would also like to know how much the average local relative frequency of the attribute can vary with configuration. By computing this second moment, we can determine how much the aggregate action in the economy and the accompanying economic multiplier vary with configuration for a given policy. To compute this second moment, $\text{Var}\,\widehat{F}_{avg}(\mathbf{Z}, N, n)$, we introduce one more piece of notation. We define random variable $D_w^-(\mathbf{Z})$ whose realizations are agent weights $[\mathbf{d}_w^-(\mathbf{Z})]_i$. By introducing random variable $D_w^-(\mathbf{Z})$, we can compactly express population moments for the set of agent weights; for example, $ED_w^-(\mathbf{Z}) = \frac{1}{N}\sum_{i=1}^{N}[\mathbf{d}_w^-(\mathbf{Z})]_i = \frac{k}{N}$ and $\text{Var}\,D_w^-(\mathbf{Z}) = \frac{1}{N}\sum_{i=1}^{N}\left([\mathbf{d}_w^-(\mathbf{Z})]_i - \frac{k}{N}\right)^2$. The closed-form expression for $\text{Var}\,\widehat{F}_{avg}(\mathbf{Z}, N, n)$ is then as follows:

Lemma 2.2 $\text{Var}\,\widehat{F}_{avg}(\mathbf{Z}, N, n) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N\,\text{Var}\,D_w^-(\mathbf{Z})\right)$.

The variance of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ fundamentally depends on the variance of agents' weights. The greater the heterogeneity in agents' weights, the greater the variation in the average local relative frequency of the attribute because this quantity then depends more strongly on which subset of agents actually has the attribute's unit value.

Lemma 2.3 shows how to compute the lower and upper bounds on the support of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$. Given that n agents have received a positive shock, we compute both the lowest and the highest possible average local relative frequencies of the attribute. We later use this result to compute the lowest and highest possible aggregate actions and economic multipliers consistent with a particular policy.

Lemma 2.3 *Construct the ordered multiset* $\{\widetilde{w}_i\}_{i=1}^{N}$ *from the elements of* $\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})$ *so that* $\widetilde{w}_i \leq \widetilde{w}_{i'}$ *whenever* $i \leq i'$. *The lower and upper bounds on the support of* $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ *are respectively:*

$$\min \operatorname{supp} \widehat{F}_{avg}(\mathbf{Z}, N, n) = \sum_{i=1}^{n} \widetilde{w}_i \quad \textit{and} \quad \max \operatorname{supp} \widehat{F}_{avg}(\mathbf{Z}, N, n) = \sum_{i=N-n+1}^{N} \widetilde{w}_i.$$

We attain the lower bound on the support of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ when the n agents with the smallest weight have the attribute's unit value. Meanwhile, we attain the upper bound on the support of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ when the n agents with the largest weight have the attribute's unit value.

We would moreover like to identify those network topologies and those vectors of agent weights, $\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})$, for which $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$ is *invariant to configuration*:

Definition 2.2 $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$ *is* invariant to configuration *when* $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n) = \widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}', N, n)$ *for all configurations* $\mathbf{b}(N, n), \mathbf{b}'(N, n) \in \mathcal{B}(N, n)$, *and this property holds for all feasible* n.

When $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$ is invariant to configuration, the distribution $G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t)$ is degenerate. The particular configuration of positive shocks among agents is irrelevant; regardless of the configuration, the average local relative frequency of the attribute is the same, which makes the distributions of aggregate output and economic multipliers likewise degenerate. In the next result, we identify the necessary values for agents' weights so that $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$ is invariant to configuration:

Lemma 2.4 *$\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n) = [\mathbf{d}_{w}^{-}(\mathbf{Z})]^{T}\mathbf{b}(N, n)$ is invariant to configuration if and only if $[\mathbf{d}_{w}^{-}(\mathbf{Z})]_i = \frac{k}{N}$ for all $i \in \{1, \ldots, N\}$. When $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$ is invariant to configuration, $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n) = \frac{kn}{N}$.*

Provided that every agent has the same weight, and in particular, the same average weighted in-degree, $\widehat{F}_{avg}(\mathbf{Z}, N, n) = \frac{kn}{N}$ with probability 1. Any network $\mathcal{G}(\mathbf{Z})$ structured so that each agent has the same weighted in-degree yields this null case in which configuration is irrelevant. The behavior of the economy here only depends on k, N, and n, and not on the underlying configuration $\mathbf{b}(N, n)$.

We proceed by returning to the original general setting in which the distribution $G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t)$ is non-degenerate. In addition to presenting closed-form expressions for the statistical features of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$, we are interested in characterizing its CDF, $G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t)$. The next lemma identifies a closed-form expression that essentially allows us to draw the CDF of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ for all feasible network structures, population sizes, and number of agents being targeted by the policy. We first introduce the function $J(\mathbf{Z}, N, n, t)$:

$$J(\mathbf{Z}, N, n, t) = \Phi(t) - H_2(t)\,\phi(t)\,C_1 \sum_{i=1}^{N} \widehat{w}_i^3 - H_3(t)\,\phi(t)\left[C_2\left(\sum_{i=1}^{N} \widehat{w}_i^4 - \frac{3}{N}\right) - \frac{1}{4N}\right] - H_5(t)\,\phi(t)\,C_3\left(\sum_{i=1}^{N} \widehat{w}_i^3\right)^2,$$

where $\widehat{w}_i = \frac{[\mathbf{d}_{w}^{-}(\mathbf{Z})]i - ED_{w}^{-}(\mathbf{Z})}{\sqrt{N\,\mathrm{Var}\,D_{w}^{-}(\mathbf{Z})}}$, $C_1 = \frac{1-\frac{2n}{N}}{6\left(\frac{n}{N}\left(1-\frac{n}{N}\right)\right)^{1/2}}$, $C_2 = \frac{1-6\left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right)}{24\left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right)}$, $C_3 = \frac{\left(1-\frac{2n}{N}\right)^2}{72\left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right)}$, $\phi(t) = \Phi'(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$, and $H_i(t)\,\phi(t) = (-1)^i \frac{d^i}{dt^i}\phi(t)$. We then approximate CDF $G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}}(t)$ by the function $J(\mathbf{Z},N,n,t)$:

Lemma 2.5 *Provided that condition (c) holds,*

$$\left| G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}}(t) - J(\mathbf{Z},N,n,t) \right| < C_4 \times \sum_{i=1}^{N} |\widehat{w}_i|^5$$

for all t, where C_4 is only a function of $\frac{n}{N}$. Condition (c) is as follows:

Condition (c) (Robinson (1978)) *Given $C' > 0$, there exist $\epsilon' > 0$, $C > 0$, and $\delta > 0$ not depending on N such that, for any fixed t, the number of indices j, for which $\left|\widehat{w}_j\,\widehat{x} - t - 2\widehat{r}\pi\right| > \epsilon'$, for all $\widehat{x} \in \left(C'\left[\max_i |\,\widehat{w}_i\,|\right]^{-1}, C\left[\sum_{i=1}^{N} |\,\widehat{w}_i\,|^5\right]^{-1}\right)$ and all $\widehat{r} = 0, \pm1, \pm2, \ldots$, is greater than δN, for all N.*

Condition (c) requires that the multiset of agent weights, $\{[\mathbf{d}_w^-(\mathbf{Z})]_i\}_{i=1}^{N}$ not be clustered around two few values. Given Lemma 2.5, we can very strongly approximate the distribution, $G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t)$:

$$G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t) \approx J\left(\mathbf{Z}, N, n, \frac{t - E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}\right).$$

Note that

$$\sum_{i=1}^{N} \widehat{w}_i^3 = N^{-1/2}\,\text{Skew}\,D_w^-(\mathbf{Z}) \;\text{ and }\; \sum_{i=1}^{N} \widehat{w}_i^4 - \frac{3}{N} = N^{-1} \times \left(\text{Excess Kurtosis}\,D_w^-(\mathbf{Z})\right).$$

We can therefore re-write the approximating function $J(\mathbf{Z},N,n,t)$ in terms of the higher-order moments of $D_w^-(\mathbf{Z})$. The asymptotic expansion $J(\mathbf{Z},N,n,t)$ is to order $1/N$.

Lastly, we characterize the limiting behavior of $G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t)$ as $N \to \infty$. We define the quantity

$$\kappa_N(\epsilon') = \frac{1}{\sum_{i=1}^{N}\left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_i - \frac{k}{N}\right)^2} \sum_{\substack{j\in\{1,\ldots,N\}\text{ s.t.}\\ \left|\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_j - \frac{k}{N}\right| > \epsilon'\sigma_N}} \left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_j - \frac{k}{N}\right)^2$$

where $\sigma_N = \left(\frac{n}{N}\left(1-\frac{n}{N}\right)\sum_{i=1}^{N}\left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_i - \frac{k}{N}\right)^2\right)^{1/2}$. We make the population size, N, explicit for the $N \times 1$ vector of agent weights (i.e., we rewrite $\mathbf{d}_w^{-}(\mathbf{Z})$ as $\mathbf{d}_{N,w}^{-}(\mathbf{Z})$) because we wish to characterize the distribution of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ as N increases. Our central limit theorem-type result is the following:

Lemma 2.6 *If* $\lim_{N\to\infty}\kappa_N(\epsilon') = 0$ *for any* $\epsilon' > 0$, *then* $\lim_{N\to\infty} G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n) - \frac{kn}{N}}{\sigma_N}}(t) = \Phi(t)$ *for all real* t, *where* $\Phi(\cdot)$ *is the standard normal CDF.*

The requirement that $\lim_{N\to\infty}\kappa_N(\epsilon') = 0$ for any $\epsilon' > 0$ is a Lindeberg-type condition. When this condition holds, we informally have that $\lim_{N\to\infty} G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t) \approx \Phi\left(\frac{t-\frac{kn}{N}}{\sigma_N}\right)$. The distribution of $\widehat{F}_{avg}(\mathbf{Z}, N, n)$ is asymptotically normal, with a mean of $\frac{kn}{N}$ and a variance that collapses to zero as the population size increases provided that the Lindeberg-type condition holds. Given this set of lemmata, we can characterize in closed form the distributions of possible aggregate actions and economic multipliers for any feasible network structure, population size, and policy targeting n agents.

2.2.3 General Environment

In this subsection, we develop the general theoretical environment. Our general environment nests the three classes of networked environments that we later consider: (1) networked environments with strategic complements and strategic substitutes, (2) networked environments with coordination and anti-coordination, and (3) networked environments with production. Now, in the general environment, we have a population of N agents who are organized on the network $\mathcal{G}(\mathbf{Z}') = (\mathcal{V}(\mathbf{Z}'), \mathcal{E}(\mathbf{Z}'))$. $\mathcal{G}(\mathbf{Z}')$ is the naturally occurring network in the environment being studied; in the environment with production, for example, it is the production network. Given the theoretical environment and implemented policy, we are interested in the resulting aggregate action and the corresponding economic multiplier. To compute these two quantities, a transformed version of $\mathcal{G}(\mathbf{Z}')$ becomes the relevant network. We refer to this new network as $\mathcal{G}(\mathbf{Z})$ with weighted adjacency matrix $\mathbf{Z}$. For each of the specific environments that we later study, we explicitly identify both $\mathcal{G}(\mathbf{Z}')$ and $\mathcal{G}(\mathbf{Z})$.

In the general environment, an outside actor enacts a policy. This policy exogenously delivers a positive shock of a fixed magnitude to a subset of $n \leq N$ agents. We introduce the $N \times 1$ vector $\boldsymbol{\rho}$ to capture the policy-induced shock. If the policy is financed by transfers, we set $[\boldsymbol{\rho}]_i = \epsilon > 0$ if agent i is receiving the positive shock, and otherwise we set $[\boldsymbol{\rho}]_i = -\frac{n\epsilon}{N-n}$. Note that $\mathbf{1}^T\boldsymbol{\rho} = 0$ when the policy is financed by transfers. Meanwhile, if the policy is financed by stimulus, we set $[\boldsymbol{\rho}]_i = \epsilon$ if agent i is receiving the positive shock, and otherwise we set $[\boldsymbol{\rho}]_i = 0$. How the shock exactly impacts agent decision-making behavior depends on the specific environment. Given $\boldsymbol{\rho}$, we can construct the configuration vector $\mathbf{b}(N,n)$: $[\mathbf{b}(N,n)]_i = 1$ if $[\boldsymbol{\rho}]_i = \epsilon$ and otherwise $[\mathbf{b}(N,n)]_i = 0$ with $\mathbf{1}^T\mathbf{b}(N,n) = n$.

Given that N agents are organized on the network $\mathcal{G}(\mathbf{Z})$ and the outside actor has enacted a policy providing a positive shock to n agents, we are interested in the resulting aggregate action and economic multiplier. In the general environment, the configuration-specific aggregate action, $y_{agg}(\mathbf{Z}, \mathbf{b}, N, n, \ell)$, takes the following form:

$$y_{agg}(\mathbf{Z}, \mathbf{b}, N, n, \ell) = y_{agg}^{no} + \gamma_1 N \left[\mathbf{d}_w^-(\mathbf{Z})\right]^T \boldsymbol{\rho}; \tag{2.1}$$

y_{agg}^{no} is the aggregate action for all agents in the population in the absence of any policy, γ_1 is an environment-specific constant, $\mathbf{1}^T \mathbf{d}_{\overline{w}}^{-}(\mathbf{Z}) = k$, and $\ell \in \{0,1\}$ is an argument that indicates whether the enacted policy is being financed by transfers or external stimulus. $\ell = 0$ denotes a policy financed by transfers and $\ell = 1$ denotes a policy financed by stimulus; the elements of vector $\boldsymbol{\rho}$ adjust depending on the particular configuration and whether $\ell = 0$ or $\ell = 1$. We compute the corresponding configuration-specific economic multiplier as follows: $m(\mathbf{Z}, \mathbf{b}, N, n, \ell) = \frac{dy_{agg}(\mathbf{Z},\mathbf{b},N,n,\ell)}{d\epsilon}$. The economic multiplier captures the change in the aggregate action given that a particular configuration of n agents is receiving ϵ units of a positive shock.

The aggregate action is, in essence, a weighted sum of shocks; the relevant weights here are agents' average weighted in-degrees, $\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})$, for the network $\mathcal{G}(\mathbf{Z})$. The higher an agent's weight, as determined by the structure of $\mathbf{Z}$, the greater the effect that the agent has on the aggregate action if he or she is the recipient of a positive shock. In general, in networked environments, the interdependencies of agents' actions can be complicated; the network $\mathcal{G}(\mathbf{Z})$ disentangles these complexities. It transforms the original network $\mathcal{G}(\mathbf{Z}')$ into a new structure $\mathcal{G}(\mathbf{Z})$ whose average weighted in-degrees simply determine the effect that each agent's action has on the aggregate action.

We now examine how to express the aggregate action when we have a particular configuration of positive shocks. Let's suppose that agents $1, \ldots, n$ are the recipients of a positive shock. We then have $b_i = 1$ for $i \in \{1, \ldots, n\}$ and otherwise $b_i = 0$. If the policy is financed by transfers, the aggregate action is as follows:

$$y_{agg}(\mathbf{Z}, \mathbf{b}, N, n, 0) = y_{agg}^{no} + \gamma_1 N \epsilon \Big[\big([\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})]_1 + \cdots + [\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})]_n \big) - \frac{n}{N-n} \Big([\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})]_{n+1} + \cdots + [\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})]_N \Big) \Big].$$

If the policy is instead financed by stimulus, the aggregate action is as follows:

$$y_{agg}\left(\mathbf{Z},\mathbf{b},N,n,1\right)=y_{agg}^{no}+\gamma_1 N\epsilon\left(\left[\mathbf{d}_w^-\left(\mathbf{Z}\right)\right]_1+\cdots+\left[\mathbf{d}_w^-\left(\mathbf{Z}\right)\right]_n\right).$$

These expressions for $y_{agg}\left(\mathbf{Z},\mathbf{b},N,n,0\right)$ and $y_{agg}\left(\mathbf{Z},\mathbf{b},N,n,1\right)$ immediately follow from Equation 2.1 after substituting in the appropriate vector $\boldsymbol{\rho}$.

Now that we have studied the aggregate action for a particular configuration of positive shocks, we introduce the accompanying random variable. We define random variable $Y_{agg}\left(\mathbf{Z},N,n,0\right)$ as the aggregate action in a setting with transfers and random variable $Y_{agg}\left(\mathbf{Z},N,n,1\right)$ as the aggregate action in a setting with stimulus. The realizations of these random variables are configuration-specific realizations of the aggregate action when a certain subset of the population receives the positive shock. The accompanying CDFs are $G_{Y_{agg}(\mathbf{Z},N,n,0)}\left(t\right)$ and $G_{Y_{agg}(\mathbf{Z},N,n,1)}\left(t\right)$, and in constructing these CDFs, we assume that every configuration is equally likely. We also define the random variables for the corresponding economic multipliers. Random variable $M\left(\mathbf{Z},N,n,0\right)=\frac{dY_{agg}(\mathbf{Z},N,n,0)}{d\epsilon}$ is the economic multiplier in a setting with transfers and random variable $M\left(\mathbf{Z},N,n,1\right)=\frac{dY_{agg}(\mathbf{Z},N,n,1)}{d\epsilon}$ is the economic multiplier in a setting with stimulus. The accompanying CDFs are $G_{M(\mathbf{Z},N,n,0)}\left(t\right)$ and $G_{M(\mathbf{Z},N,n,1)}\left(t\right)$; in constructing these CDFs, we assume as well that each configuration is equally likely.

Given Equation 2.1, we can write the random variables for aggregate output and the corresponding economic multiplier in a setting with transfers as follows:

$$Y_{agg}\left(\mathbf{Z},N,n,0\right)=y_{agg}^{no}+\gamma_1\frac{N^2\epsilon}{N-n}\times\left(\widehat{F}_{avg}\left(\mathbf{Z},N,n\right)-\frac{kn}{N}\right)\text{ and}\tag{2.2}$$

$$M\left(\mathbf{Z},N,n,0\right)=\gamma_1\frac{N^2}{N-n}\times\left(\widehat{F}_{avg}\left(\mathbf{Z},N,n\right)-\frac{kn}{N}\right).\tag{2.3}$$

Meanwhile, the random variables for aggregate output and the corresponding economic multiplier in a setting with externally funded stimulus are as follows:

$$Y_{agg}(\mathbf{Z}, N, n, 1) = y_{agg}^{no} + \gamma_1 N\epsilon \times \widehat{F}_{avg}(\mathbf{Z}, N, n) \text{ and} \tag{2.4}$$

$$M(\mathbf{Z}, N, n, 1) = \gamma_1 N \times \widehat{F}_{avg}(\mathbf{Z}, N, n). \tag{2.5}$$

Depending on the particular configuration of the positive shock among agents in the population, we can have variation in its average local relative frequency. The higher the average local relative frequency of the positive shock, the higher the aggregate action and the higher the economic multiplier.

We proceed by building on the set of lemmata from the previous subsection. Here, we present theoretical results in which we characterize the distributional features of the aggregate action and the corresponding economic multiplier in settings with transfers and in settings with stimulus. We first compute in closed form the first moment for the distributions of possible aggregate actions and economic multipliers:

Proposition 2.1 *The first moments for the aggregate action and economic multiplier are:*

$$EY_{agg}(\mathbf{Z}, N, n, 0) = y_{agg}^{no}, \qquad EM(\mathbf{Z}, N, n, 0) = 0,$$

$$EY_{agg}(\mathbf{Z}, N, n, 1) = y_{agg}^{no} + \gamma_1 kn\epsilon, \quad \textit{and} \quad EM(\mathbf{Z}, N, n, 1) = \gamma_1 kn.$$

In settings with transfers, the mean aggregate action is equal to its no-intervention level, y_{agg}^{no}, and the mean economic multiplier is equal to zero. When the outside actor finances its positive shock to a subset of agents by requesting a transfer of funds from the other agents in the population, the policy has no aggregate effect on average. This result holds for every feasible agent interaction structure, $\mathcal{G}(\mathbf{Z})$, population size, N, and number of agents being targeted for a positive shock, n. When the policy is instead financed by stimulus, the mean value of the aggregate action can deviate from its no-intervention level and the corresponding mean economic multiplier can also deviate from zero. The values of these first moments in settings with stimulus depend both on k, which we derive from $\mathbf{Z}$, and γ_1 which depends on the particular networked environment. When the network and environment are such that $k = 1$ and $\gamma_1 = 1$, then the increase in the mean aggregate action above its no-intervention level is equal to the aggregate amount of stimulus, $n\epsilon$.

The next proposition computes in closed form the second moment for the distributions of aggregate actions and economic multipliers:

Proposition 2.2 *The second moments for the aggregate action and economic multiplier are:*

$$\begin{aligned}
\operatorname{Var} Y_{agg}(\mathbf{Z}, N, n, 0) &= \left(\gamma_1 \frac{N^2\epsilon}{N-n}\right)^2 \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} D_w^-(\mathbf{Z})\right), \\
\operatorname{Var} M(\mathbf{Z}, N, n, 0) &= \left(\gamma_1 \frac{N^2}{N-n}\right)^2 \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} D_w^-(\mathbf{Z})\right), \\
\operatorname{Var} Y_{agg}(\mathbf{Z}, N, n, 1) &= (\gamma_1 N\epsilon)^2 \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} D_w^-(\mathbf{Z})\right), \textit{ and} \\
\operatorname{Var} M(\mathbf{Z}, N, n, 1) &= (\gamma_1 N)^2 \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} D_w^-(\mathbf{Z})\right).
\end{aligned}$$

These second moments depend on the environment-specific constant, γ_1, the fraction of agents receiving a positive shock, $\frac{n}{N}$, the total population size, N, and the distribution of agents' weights, $G_{D_w^-(\mathbf{Z})}(t)$. Each agent's weight determines that agent's effect on the aggregate action if he or she is the recipient of a positive shock or, in the setting with transfers, the recipient of a negative shock. When there is a large amount of heterogeneity in agents' weights, the overall effect on the aggregate action strongly varies depending on which configuration of agents is receiving a positive shock. Therefore, the variance of the aggregate action and the variance of the economic multiplier directly depend in the variance of agents' weights.

In Proposition 2.3, we compute the lowest and the highest possible aggregate actions and the lowest and the highest possible economic multipliers given that n agents in a total population of N agents are receiving a positive shock:

Proposition 2.3 *Construct the ordered multiset* $\{\widetilde{w}_i\}_{i=1}^{N}$ *from the elements of* $\mathbf{d}_w^-(\mathbf{Z})$ *so that* $\widetilde{w}_i \leq \widetilde{w}_{i'}$ *whenever* $i \leq i'$*. Given Equations 2.2-2.5, we compute* $\min\operatorname{supp} Y_{agg}(\mathbf{Z}, N, n, 0)$, $\min\operatorname{supp} M(\mathbf{Z}, N, n, 0)$, $\min\operatorname{supp} Y_{agg}(\mathbf{Z}, N, n, 1)$*, and* $\min\operatorname{supp} M(\mathbf{Z}, N, n, 1)$ *by setting* $\widehat{F}_{avg}(\mathbf{Z}, N, n) = \sum_{i=1}^{n} \widetilde{w}_i$*, and we compute* $\max\operatorname{supp} Y_{agg}(\mathbf{Z}, N, n, 0)$*,* $\max\operatorname{supp} M(\mathbf{Z}, N, n, 0)$*,* $\max\operatorname{supp} Y_{agg}(\mathbf{Z}, N, n, 1)$*, and* $\max\operatorname{supp} M(\mathbf{Z}, N, n, 1)$ *by setting* $\widehat{F}_{avg}(\mathbf{Z}, N, n) = \sum_{i=N-n+1}^{N} \widetilde{w}_i$*.*

The lower and upper bounds on the support of the aggregate action and the economic multiplier directly depend on the topology of the network $\mathcal{G}(\mathbf{Z})$. We compute the lower bound by supposing that the n agents with the smallest average weighted in-degrees receive a positive shock. Meanwhile, we compute the upper bound by supposing that the n agents with the largest average weighted in-degrees receive a positive shock. The extent to which the lower and upper bounds on the support of the aggregate action and the support of the economic multiplier differ from each other depend on the extent to which the smallest and largest average weighted in-degrees for the graph $\mathcal{G}(\mathbf{Z})$ differ from each other.

When we have heterogeneity in agents' weights, $\operatorname{Var} D_w^-(\mathbf{Z})$ is non-zero, which makes the distributions of possible aggregate actions and economic multipliers also have positive variance. We refer to these distributions of aggregate actions and economic multipliers as being non-degenerate. Depending on the particular configuration of positive shocks, we experience variation in both the aggregate action and the economic multiplier. The next result focuses on the null case, in which the distributions of aggregate actions and economic multipliers are degenerate; regardless of the particular configuration of positive shocks among agents in the population, holding n fixed, both the aggregate action and the economic multiplier remain unchanged. This result identifies the necessary condition for degeneracy and the resulting values for the aggregate action and economic multiplier:

Proposition 2.4 *$y_{agg}(\mathbf{Z}, \mathbf{b}, N, n, 0)$, $m(\mathbf{Z}, \mathbf{b}, N, n, 0)$, $y_{agg}(\mathbf{Z}, \mathbf{b}, N, n, 1)$, and $m(\mathbf{Z}, \mathbf{b}, N, n, 1)$ are all invariant to configuration if and only if $[\mathbf{d}_w^-(\mathbf{Z})]_i = \frac{k}{N}$ for all $i \in \{1, \ldots, N\}$. When these four quantities are invariant to configuration, $Y_{agg}(\mathbf{Z}, N, n, 0) = y_{agg}^{no}$, $M(\mathbf{Z}, N, n, 0) = 0$, $Y_{agg}(\mathbf{Z}, N, n, 1) = y_{agg}^{no} + \gamma_1 kn\epsilon$, and $M(\mathbf{Z}, N, n, 1) = \gamma_1 kn$, all with probability 1.*

When every agent has the same weight, that is, when $[\mathbf{d}_w^-(\mathbf{Z})]_i = \frac{k}{N}$ for all $i \in \{1, \ldots, N\}$, the distributions of the aggregate action and economic multiplier are degenerate for all values n. In a setting with transfers, the aggregate action is equal to its no-intervention level and the economic multiplier is equal to zero with probability 1. Therefore, any policy that the outside actor implements via transfer is ineffective; there is no possibility for adjustment to the aggregate action. However, in a setting with stimulus, the aggregate action can deviate from its no-intervention level and the economic multiplier can deviate from zero; the effect of the policy on both the aggregate action and the economic multiplier in a setting with stimulus indeed does depend on the values of k and γ_1, but regardless of the configuration, holding n fixed, the effect is always the same.

We return to the original setting in which the distributions of possible aggregate actions and economic multipliers are non-degenerate. Thus far, we have been presenting results that characterize certain features of these distributions, namely, their first and second moments and the bounds on their supports. The next result shows us how to actually draw the CDFs for these distributions:

Proposition 2.5 *For $\ell \in \{0,1\}$, provided that condition (c) of Lemma 2.5 holds,*

$$\left| G_{\frac{Y_{agg}(\mathbf{Z},N,n,\ell)-EY_{agg}(\mathbf{Z},N,n,\ell)}{\left(\mathrm{Var}\,Y_{agg}(\mathbf{Z},N,n,\ell)\right)^{1/2}}}(t) - J(\mathbf{Z},N,n,t) \right| < C_4 \times \sum_{i=1}^{N} |\widehat{w}_i|^5 \text{ and}$$

$$\left| G_{\frac{M(\mathbf{Z},N,n,\ell)-EM(\mathbf{Z},N,n,\ell)}{\left(\mathrm{Var}\,M(\mathbf{Z},N,n,\ell)\right)^{1/2}}}(t) - J(\mathbf{Z},N,n,t) \right| < C_4 \times \sum_{i=1}^{N} |\widehat{w}_i|^5$$

for all t, where C_4 is only a function of $\frac{n}{N}$ and $\widehat{w}_i = \frac{\left[\mathbf{d}_w^-(\mathbf{Z})\right]_i - ED_w^-(\mathbf{Z})}{\sqrt{N\,\mathrm{Var}\,D_w^-(\mathbf{Z})}}$.

Given this result, for $\ell \in \{0,1\}$, we can strongly approximate the CDFs for the aggregate action and the corresponding economic multiplier as follows:

$$G_{Y_{agg}(\mathbf{Z},N,n,\ell)}(t) \approx J\left(\mathbf{Z},N,n,\frac{t-EY_{agg}(\mathbf{Z},N,n,\ell)}{\left(\mathrm{Var}\,Y_{agg}(\mathbf{Z},N,n,\ell)\right)^{1/2}}\right) \text{ and}$$

$$G_{M(\mathbf{Z},N,n,\ell)}(t) \approx J\left(\mathbf{Z},N,n,\frac{t-EM(\mathbf{Z},N,n,\ell)}{\left(\mathrm{Var}\,M(\mathbf{Z},N,n,\ell)\right)^{1/2}}\right).$$

We observe that these approximations depend on the function $J(\mathbf{Z},N,n,t)$. The function $J(\mathbf{Z},N,n,t)$ is an asymptotic expansion whose first term is the normal distribution and whose other terms represent deviations away from the normal distribution. The extent to which these other terms are non-zero depends on the extent to which the distribution of agent weights, $G_{D_w^-(\mathbf{Z})}(t)$, has non-zero skewness and/or non-zero excess kurtosis. When the distribution of agent weights has these non-zero higher-order moments, then the CDFs $G_{Y_{agg}(\mathbf{Z},N,n,\ell)}(t)$ and $G_{M(\mathbf{Z},N,n,\ell)}(t)$, for $\ell \in \{0,1\}$, deviate from distributions that are normal. It is ultimately the topology of $\mathcal{G}(\mathbf{Z})$ that shapes the lower-order and higher-order distributional features of $G_{Y_{agg}(\mathbf{Z},N,n,\ell)}(t)$ and $G_{M(\mathbf{Z},N,n,\ell)}(t)$, for $\ell \in \{0,1\}$. The topology of the network can generate distributions with properties of skewness and/or heavy-tailedness based on the statistical features of its accompanying network-derived vector of agent weights, $\mathbf{d}_w^-(\mathbf{Z})$. Given Proposition 2.5, we can draw the CDFs of aggregate actions and economic multipliers for any feasible network structure, $\mathcal{G}(\mathbf{Z})$, population size, N, and number of agents receiving a positive shock, n. For any policy, we can draw, via closed-form expressions, the resulting distribution of possible aggregate actions and the resulting distribution of possible economic multipliers.

We proceed to characterize limiting distributions for the aggregate action and economic multiplier as $N \to \infty$:

Proposition 2.6 *If* $\lim_{N\to\infty} \kappa_N(\epsilon') = 0$ *for any* $\epsilon' > 0$, *then*
$\lim_{N\to\infty} G_{\frac{Y_{agg}(\mathbf{Z},N,n,\ell)-EY_{agg}(\mathbf{Z},N,n,\ell)}{(\text{Var}\, Y_{agg}(\mathbf{Z},N,n,\ell))^{1/2}}}(t) = \Phi(t)$ *and* $\lim_{N\to\infty} G_{\frac{M(\mathbf{Z},N,n,\ell)-EM(\mathbf{Z},N,n,\ell)}{(\text{Var}\, M(\mathbf{Z},N,n,\ell))^{1/2}}}(t) = \Phi(t)$
for $\ell \in \{0,1\}$ *and for all real* t.

When the Lindeberg-type condition is satisfied, as $N \to \infty$, the aggregate action and economic multiplier become normally distributed. Informally,
$\lim_{N\to\infty} G_{Y_{agg}(\mathbf{Z},N,n,\ell)}(t) \approx \Phi\left(\frac{t-EY_{agg}(\mathbf{Z},N,n,\ell)}{(\text{Var}\, Y_{agg}(\mathbf{Z},N,n,\ell))^{1/2}}\right)$, and $\lim_{N\to\infty} G_{M(\mathbf{Z},N,n,\ell)}(t) \approx \Phi\left(\frac{t-EM(\mathbf{Z},N,n,\ell)}{(\text{Var}\, M(\mathbf{Z},N,n,\ell))^{1/2}}\right)$. Even though we are studying the limiting case in which $N \to \infty$, in this particular setting, $\text{Var}\, Y_{agg}(\mathbf{Z},N,n,\ell)$ and $\text{Var}\, M(\mathbf{Z},N,n,\ell)$ for $\ell \in \{0,1\}$ do not generally tend to zero. To see this, let us examine $\text{Var}\, Y_{agg}(\mathbf{Z},N,n,0)$. From Proposition 2.2,

$$\text{Var}\, Y_{agg}(\mathbf{Z},N,n,0) = \left(\gamma_1 \frac{N}{N-n} N\epsilon\right)^2 \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}\sum_{i=1}^{N}\left(\left[\mathbf{d}_w^-(\mathbf{Z})\right]_i - \frac{k}{N}\right)^2.$$

Let us hold $\frac{n}{N}$ fixed as N grows. Then:

$$\text{Var}\, Y_{agg}(\mathbf{Z},N,n,0) \propto N^2\left[\sum_{i=1}^{N}\left(\left[\mathbf{d}_w^-(\mathbf{Z})\right]_i - \frac{k}{N}\right)^2\right],$$

so the behavior of $\text{Var}\, Y_{agg}(\mathbf{Z},N,n,0)$ depends on how $\mathbf{d}_w^-(\mathbf{Z})$ and $\frac{k}{N}$ evolve as the population grows; we can imagine that there are many scenarios in which $\text{Var}\, Y_{agg}(\mathbf{Z},N,n,0)$ either does not tend toward zero or tends to zero very slowly. As a result, the particular configuration of positive shocks remains relevant for all population sizes; even for large N, there will still be variation in both the aggregate action and the economic multiplier across configurations.

Care must be taken in determining whether the Lindeberg-type condition actually gets satisfied, that is, $\lim_{N\to\infty} \kappa_N(\epsilon') = 0$ for any $\epsilon' > 0$. When $\mathbf{Z}$ is row-stochastic, the Lindeberg-type condition is generally satisfied. We have $k = 1$ for all population sizes N; moreover, agents' weights, $[\mathbf{d}_w^-(\mathbf{Z})]_i$, are non-negative and constrained to sum to 1, so as N increases, agents' weights generally tend toward zero, and they become increasingly closer to the average agent weight. However, there exist many types of matrices $\mathbf{Z}$ for which the Lindeberg-type condition does not get satisfied. We can imagine that there exist classes of matrices $\mathbf{Z}$ for which k changes as N grows; for example, there exist growing matrices $\mathbf{Z}$ for which $\frac{k}{N} = 1$ for all N. Then, in such settings, agents' weights, $[\mathbf{d}_w^-(\mathbf{Z})]_i$, need not move closer to the average agent weight, $\frac{k}{N}$, as $N \to \infty$.

In Proposition 2.7, we compute in closed form the probability that a policy targeting n agents leads to an aggregate action below its no-intervention level and a negative multiplier:

Proposition 2.7 *The probability of an aggregate action below its no-intervention level and a negative multiplier are as follows:*

$$\Pr\left[Y_{agg}(\mathbf{Z}, N, n, 0) < y_{agg}^{no}\right] = \Pr\left[M(\mathbf{Z}, N, n, 0) < 0\right] = \Pr\left[\widehat{F}_{avg}(\mathbf{Z}, N, n) < \frac{kn}{N}\right],$$

and

$$\Pr\left[Y_{agg}(\mathbf{Z}, N, n, 1) < y_{agg}^{no}\right] = \Pr\left[M(\mathbf{Z}, N, n, 1) < 0\right] = \Pr\left[\widehat{F}_{avg}(\mathbf{Z}, N, n) < 0\right].$$

Provided that condition (c) of Lemma 2.5 holds, $\Pr\left[\widehat{F}_{avg}(\mathbf{Z}, N, n) < \frac{kn}{N}\right] \approx J(\mathbf{Z}, N, n, 0)$ *and* $\Pr\left[\widehat{F}_{avg}(\mathbf{Z}, N, n) < 0\right] \approx J\left(\mathbf{Z}, N, n, -\frac{E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\text{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}\right)$*, with*

$\widehat{w}_i = \frac{\left[\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z})\right]_i - ED_{\overline{w}}^{-}(\mathbf{Z})}{\sqrt{N \operatorname{Var} D_{\overline{w}}^{-}(\mathbf{Z})}}.$

Provided that condition (c) of Lemma 2.5 holds, we can compute in closed form the probability that a policy targeting n agents lowers the aggregate action and generates a negative economic multiplier. We can compute this probability in both settings with transfers and settings with stimulus. We can moreover compute this probability for any feasible network structure, $\mathcal{G}(\mathbf{Z})$, population size, N, and number of agents, n, being targeted by the outside actor's policy. When the outside actor's policy is financed by transfers and the network structure, $\mathcal{G}(\mathbf{Z})$, is such that the distribution of agent weights, $G_{D_{\overline{w}}^{-}(\mathbf{Z})}(t)$, has zero skewness and zero excess kurtosis, $J(\mathbf{Z}, N, n, 0) = 0.50$; the probability that the policy targeting n agents leads to a negative multiplier and a reduction in the aggregate action below its no-intervention level is equal to 50 percent. The topological features of the network $\mathcal{G}(\mathbf{Z})$ shape the distributional features of $D_{\overline{w}}^{-}(\mathbf{Z})$ and thereby determine the probability that a policy targeting n agents generates a reduction in the aggregate action and a negative economic multiplier.

Negative multipliers emerge in settings with transfers when agent weights are not all equal:

Proposition 2.8 *For every $n \in \{1, \ldots, N-1\}$, provided that $\mathbf{d}_{\overline{w}}^{-}(\mathbf{Z}) \neq \frac{k}{N}\mathbf{1}$,*
$\Pr\left[Y_{agg}(\mathbf{Z}, N, n, 0) < y_{agg}^{no}\right] = \Pr\left[M(\mathbf{Z}, N, n, 0) < 0\right] > 0.$

Given any policy that targets n agents and is financed by transfers, there is a positive probability of a negative multiplier and a positive probability that the aggregate action can be less than its no-intervention level. Practically every network structure $\mathcal{G}(\mathbf{Z})$ generates negative economic multipliers in settings with transfers. The only class of networks for which there is zero probability of a negative multiplier is the one for which $\mathbf{1}^T\mathbf{Z} = k\mathbf{1}^T$. For this particular class, the configuration of positive shocks is irrelevant, so $M(\mathbf{Z}, N, n, 0) = 0$ with probability 1. Once configuration becomes relevant, negative economic multipliers naturally emerge.

Now that we have finished characterizing policy-induced distributions of possible aggregate actions and economic multipliers in a general networked environment, we transition towards studying policy-induced distributions of aggregate actions and economic multipliers in three specific networked environments.

2.3 Networked Environments with Strategic Complements and Strategic Substitutes

We transition to our first setting of network-based interaction among agents. The environment that we study in this section is one of strategic complementarities and strategic substitutabilities. In this environment, the action of an agent can potentially tilt away from its autarkic level depending on the network of linkages and the extent to which other agents' actions act as complements or substitutes. We are able to characterize the distribution of possible aggregate actions and the distribution of possible multipliers in such a setting when a random subset of networked agents in the economy receives either a transfer of wealth or stimulus. Our model builds on work by Ballester, Calvó-Armengol, and Zenou (2006), which studies network games with linear-quadratic payoffs.

Let's consider an economy with a population of N agents. Each agent $i \in \{1, \ldots, N\}$ chooses an action $y_i \geq 0$ and receives a payoff

$$u_i(y_1, \ldots, y_N) = \alpha_i y_i + \frac{1}{2}\sigma y_i^2 + \sum_{\substack{j=1 \\ j \neq i}}^{N} \sigma_{ij} y_i y_j,$$

with $\alpha_i > 0$. Agents' payoffs are strictly concave in their own individual actions:

$\frac{\partial^2 u_i}{\partial y_i^2} = \sigma < 0$. Bilateral influences on agent i's payoff are the quantities: $\frac{\partial^2 u_i}{\partial y_i \partial y_j} = \sigma_{ij}$. When $\sigma_{ij} > 0$, agent j's action is a strategic complement to agent i's action. When $\sigma_{ij} < 0$, agent j's action is a strategic substitute to agent i's action. In the absence of bilateral influences, agent i's autarkic action is $y_i^* = -\frac{\alpha_i}{\sigma}$. Actions that are strategic complements push agent i's action above its autarkic level, while actions that are strategic substitutes push agent i's action below its autarkic level. We capture the interdependencies of agents' behavior with the matrix $\mathbf{Z}' = \mathbf{\Sigma}$ whose diagonal elements are σ and off-diagonal elements are σ_{ij}. $\mathbf{\Sigma}$ is the weighted adjacency matrix that corresponds to network $\mathcal{G}(\mathbf{\Sigma})$. Parameter α_i, the marginal benefit accrued from an additional unit of action by agent i, can vary across individuals. Its value depends on each agent's wealth, that is, $\alpha_i = \psi\omega_i$ for $\psi > 0$ and wealth ω_i. Greater wealth increases agent i's action. The optimization problem for each agent i is therefore:

$$\max_{y_i} \; \psi\omega_i y_i + \frac{1}{2}\sigma y_i^2 + \sum_{\substack{j=1 \\ j \neq i}}^{N} \sigma_{ij} y_i y_j.$$

We map this theoretical environment to two different real-world settings. The first setting concerns a population of students and the amount of effort that they exert towards their education. This amount of effort depends on the behavior of their peers; see, for example, Calvó-Armengol, Patacchini, and Zenou (2009), Sacerdote (2011), and Epple and Romano (2011) for research on peer effects in education. The social network $\mathcal{G}(\mathbf{\Sigma})$ captures these peer effects. In this model, the amount of effort also depends on the wealth of the student's family. Björklund and Salvanes (2011) documents the positive relationship between family income and educational attainment; we might imagine that student effort strongly correlates with educational attainment. Family wealth raises the marginal utility of a student's effort through a variety of channels: for example, the family can afford to live in a neighborhood with higher-quality public schools that better motivate students to perform; the family can pay for enrichment activities that make learning more exciting and therefore more rewarding for the student; and the family can pay for tutoring, which increases the return that the student receives on every unit of effort. Here, the aggregate action is the aggregate effort of all students; greater aggregate effort is positively correlated with greater aggregate earning potential.

The second setting concerns the R&D divisions of firms and the amount of effort that they each allocate towards innovation. Among different R&D divisions, there is an underlying network of collaborators and competitors (Goyal and Moraga-Gonzalez (2001), König, Liu, and Zenou (2018)). The network $\mathcal{G}(\mathbf{\Sigma})$ captures these relationships as well as the extent to which firms' R&D efforts acts as strategic complements or strategic substitutes to each other. Within formal R&D alliances, knowledge spillovers can boost the productivity for the R&D divisions of linked firms and thereby serve as a strategic complement; alternatively, they can reduce the incentives for a linked firm to engage in R&D activity (see D'Aspremont and Jacquemin (1988) and Suzumura (1992)). Meanwhile, R&D activity by competitors in a firm's product space can spark either positive or negative adjustments to the amount of effort that a firm exerts for its own R&D. Effort towards R&D is costly. The marginal benefit of effort for the R&D division is the reward of innovation. For every additional unit of effort, the reward of innovation depends on the likelihood of innovation. We can imagine that this likelihood of innovation scales with the firm's wealth. The greater the firm's wealth, the more productive and innovative are the employees that the firm hires, and therefore the more likely innovation will take place. Here, the aggregate action is aggregate R&D effort across all firms; greater effort generally leads to greater innovation.

We now return to the original setup of our model and define the unique, interior Nash equilibrium in this setting. Consistent with Ballester, Calvó-Armengol, and Zenou (2006), we introduce some additional notation. We set $\underline{\sigma} = \min\left\{\sigma_{ij} | i \neq j\right\}$, $\overline{\sigma} = \max\left\{\sigma_{ij} | i \neq j\right\}$, and $\gamma = -\min\{\underline{\sigma}, 0\} \geq 0$. We assume that $\sigma < \min\{\underline{\sigma}, 0\}$. We set $\lambda = \overline{\sigma} + \gamma$, which we take to be positive. We then define the zero-diagonal non-negative square matrix $\mathbf{G}$ whose off-diagonal elements are $[\mathbf{G}]_{ij} = \frac{\sigma_{ij}+\gamma}{\lambda} \in [0,1]$. Constant $\beta = -\gamma - \sigma > 0$ and $r(\mathbf{G})$ is the spectral radius of $\mathbf{G}$. Our unique equilibrium is as follows:

Proposition 2.9 *Provided that $\beta > \lambda r(\mathbf{G})$, the unique interior Nash equilibrium in pure strategies is $\mathbf{y}^* = -\psi \boldsymbol{\Sigma}^{-1} \boldsymbol{\omega}$.*

Vector $\boldsymbol{\omega}$ denotes agents' wealth prior to any transfers or receipt of stimulus. Given agents' equilibrium behavior, we can compute the aggregate action, y_{agg}^{no}, for all agents in the population in the absence of any transfers or stimulus:

$$y_{agg}^{no} = \mathbf{1}^T \mathbf{y}^* = \psi N \left[\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right)\right]^T \boldsymbol{\omega},$$

where $\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right) = \frac{1}{N} \left(-\boldsymbol{\Sigma}^{-1}\right)^T \mathbf{1}$. The aggregate action crucially depends on the structure of agents' interaction network. As the topology of agents' interaction network changes, agents' individual actions as well as the aggregate action adjust. We essentially have two relevant networks: (1) the original agent interaction structure, $\mathcal{G}(\mathbf{Z}') = \mathcal{G}(\boldsymbol{\Sigma})$, that captures strategic complementarities and substitutabilities between agents, and (2) the network, $\mathcal{G}(\mathbf{Z}) = \mathcal{G}\left(-\boldsymbol{\Sigma}^{-1}\right)$, that determines each agent's effective weight in the population. The vector that captures agents' weights is the vector of average weighted in-degrees for the graph $\mathcal{G}\left(-\boldsymbol{\Sigma}^{-1}\right)$: $\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right)$. Agents' weights sum to k: $\mathbf{1}^T \left[\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right)\right] = k$.

Agents' weights, $\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right)$, determine how much of an effect targeted stimulus or a targeted transfer has on the aggregate action. We have this complex web of interactions. A positive shock to an agent's wealth increases that agent's autarkic action. However, actions are not decided by agents in isolation. A shock to an agent's wealth adjusts the actions of his or her neighbors, which then adjusts the actions of the neighbors of that agent's original set of neighbors, etc. The vector of agent weights, $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)$, condenses all of these effects; the larger an agent's weight, the greater the effect on the aggregate action following a shock to that agent's wealth.

We can examine what happens to the aggregate action for a particular configuration of transfers or stimulus, holding fixed agents' interaction structure. The configuration vector $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ identifies which subset of $n \leq N$ agents is receiving a positive adjustment to wealth. Element $b_i = 1$ if agent i is receiving a positive transfer or stimulus, and otherwise $b_i = 0$. Agents' wealth following either a transfer or stimulus changes from $\boldsymbol{\omega}$ to $\boldsymbol{\omega} + \boldsymbol{\rho}$. In a setting with transfers, $[\boldsymbol{\rho}]_i = \epsilon$ if $b_i = 1$ and $[\boldsymbol{\rho}]_i = -\frac{n\epsilon}{N-n}$ if $b_i = 0$. In a setting with stimulus, $[\boldsymbol{\rho}]_i = \epsilon$ if $b_i = 1$ and $[\boldsymbol{\rho}]_i = 0$ if $b_i = 0$.

We can characterize the aggregate action and the economic multiplier on the aggregate action when agents $1, \ldots, n$ receive a positive transfer of wealth and agents $n+1, \ldots, N$ receive a negative transfer of wealth so that there is a zero net transfer. In this setting, $b_i = 1$ for $i \in \{1, \ldots, n\}$:

$$\begin{aligned} y_{agg}\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 0\right) = y_{agg}^{no} + \psi N \epsilon \Bigg[\bigg(\left[\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)\right]_1 + \cdots + \left[\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)\right]_n \bigg) \\ - \frac{n}{N-n} \bigg(\left[\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)\right]_{n+1} + \cdots + \left[\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)\right]_N \bigg) \Bigg] \end{aligned}$$

The fifth argument of $y_{agg}\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 0\right)$, that is, the 0, denotes the setting in which there is a transfer of wealth. The multiplier is $m\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 0\right) = \frac{dy_{agg}\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 0\right)}{d\epsilon}$. We can also characterize the aggregate action and the economic multiplier on the aggregate action when agents $1, \ldots, n$ receive positive stimulus while agents $n+1,\ldots,N$ receive zero adjustment to wealth. In this setting with stimulus, $b_i = 1$ for $i \in \{1,\ldots,n\}$:

$$y_{agg}\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 1\right) = y_{agg}^{no} + \psi N \epsilon \left(\left[\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)\right]_1 + \cdots + \left[\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)\right]_n\right).$$

The fifth argument of $y_{agg}\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 1\right)$, that is, the 1, denotes the setting in which stimulus is externally funded. The multiplier is $m\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 1\right) = \frac{dy_{agg}\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n, 1\right)}{d\epsilon}$.

We have computed the aggregate action and the corresponding economic multiplier given a particular configuration of transfers and a particular configuration of stimulus. We are interested in all possible aggregate actions and all possible economic multipliers when n agents each receive $\epsilon > 0$ units of additional wealth. We therefore introduce random variables that allow us to characterize the distribution of possible aggregate actions as well as the distribution of possible economic multipliers given n:

Proposition 2.10 *In a setting with transfers, the aggregate action and the corresponding economic multiplier are:*

$$Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = y_{agg}^{no} + \frac{\psi N^2 \epsilon}{N-n}\left[\widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) - \frac{kn}{N}\right] \text{ and}$$
$$M\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = \frac{\psi N^2}{N-n}\left[\widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) - \frac{kn}{N}\right].$$

In a setting with stimulus, the aggregate action and the corresponding economic multiplier are:

$$Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) = y_{agg}^{no} + \psi N \epsilon \widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) \text{ and}$$
$$M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) = \psi N \widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right).$$

The expressions for $Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right)$, $M\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right)$, $Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$, and $M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ in Proposition 2.10 directly map to Equations 2.2-2.5 in Section 2.2, where we set $\mathbf{Z} = -\mathbf{\Sigma}^{-1}$ and $\gamma_1 = \psi$.

Therefore, all of the results from Section 2.2 that characterize the distributions of aggregate actions and economic multipliers map to the present setting. For example, we can analytically compute the moments of these distributions. Given the topology of agents' interaction network, we can identify the lowest possible economic multiplier and the highest possible economic multiplier on the aggregate action that is consistent with a particular fraction of the population receiving positive wealth transfers or externally financed stimulus. Moreover, we can analytically determine the probability that a wealth transfer leads to a negative economic multiplier and the probability that externally financed stimulus leads to a negative economic multiplier; when the multiplier is negative, we have a reduction in the aggregate action below the level y_{agg}^{no}.

There do indeed exist interaction structures $\mathbf{\Sigma}$ for which the multiplier on agents' aggregate action is negative, even when every agent receives a non-negative shock to wealth. In such settings, negative multipliers emerge from the strategic substitutability of agents' actions. In general, we need there to be some level of strategic substitutability to generate a negative economic multiplier:

Proposition 2.11 *In environments without strategic substitutes, provided that* $-\sigma > r\left(\mathbf{\Sigma} - \sigma \mathbf{I}\right)$, $\Pr\left[Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) \geq y_{agg}^{no}\right] = 1$ *and* $\Pr\left[M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) \geq 0\right] = 1$.

The requirement that $-\sigma > r\left(\mathbf{\Sigma} - \sigma \mathbf{I}\right)$ ensures the non-singularity of $\mathbf{\Sigma}$ and hence its invertibility: $\mathbf{\Sigma}^{-1}$ exists and we can construct $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)$. As demonstrated in the proof for Proposition 2.11, provided that $-\sigma > r\left(\mathbf{\Sigma} - \sigma \mathbf{I}\right)$, every agent has a non-negative weight in environments without strategic substitutes: $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) \geq \mathbf{0}$. From the expressions for $Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ and $M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ in Proposition 2.10, we know that any configuration of stimulus leads to a weak increase in the aggregate action and a non-negative economic multiplier whenever $\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right) \geq \mathbf{0}$. Therefore, in environments without strategic substitutes, we can never have a reduction in the aggregate action following stimulus. However, in settings with transfers, there is generally still a positive probability of negative economic multipliers, even in environments with only strategic complements. Negative multipliers emerge in such settings due to the mixture of positive and negative shocks to wealth that agents are experiencing and the heterogeneity in agents' effective weights, as captured by $\mathbf{d}_w^- \left(-\boldsymbol{\Sigma}^{-1}\right)$.

We would like to characterize both the aggregate action and the corresponding economic multiplier in two different null settings. For the first null setting, there is no underlying network:

Proposition 2.12 *In the absence of any network-based interaction, that is,* $\boldsymbol{\Sigma} = \sigma \mathbf{I}$*,*

$$Y_{agg}\left(-\boldsymbol{\Sigma}^{-1}, N, n, 0\right) = y_{agg}^{no} \quad \textit{and} \quad M\left(-\boldsymbol{\Sigma}^{-1}, N, n, 0\right) = 0$$

with probability 1, and

$$Y_{agg}\left(-\boldsymbol{\Sigma}^{-1}, N, n, 1\right) = y_{agg}^{no} - \frac{\psi \epsilon n}{\sigma} \quad \textit{and} \quad M\left(-\boldsymbol{\Sigma}^{-1}, N, n, 1\right) = -\frac{\psi n}{\sigma}$$

with probability 1.

When agents' actions no longer serve as strategic complements or strategic substitutes for each other, $\boldsymbol{\Sigma} = \sigma \mathbf{I}$ and each agent chooses an autarkic action. Transfers of wealth across agents have no effect on the aggregate action; the economic multiplier is zero with probability 1. Externally funded stimulus does lead to an increase in the aggregate action; note that $-\frac{\psi n}{\sigma} > 0$ because $\sigma < 0$. The more agents that receive $\epsilon > 0$ units of stimulus, the higher the aggregate action. The existence of a non-trivial network structure causes the aggregate action and the corresponding economic multiplier to deviate in either direction away from these autarkic values.

For the second null setting, the aggregate action is invariant to the particular configuration of transfers or stimulus. We identify the necessary restrictions on $\mathbf{\Sigma}$ that make the aggregate action invariant to configuration, and we then solve for the resulting values of the aggregate action and the corresponding economic multiplier:

Proposition 2.13 *If $\mathbf{1}^T\mathbf{\Sigma} = \delta\mathbf{1}^T$ for some $\delta \in \mathbb{R}$, both the aggregate action and the economic multiplier are invariant to configuration:*

$$Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = y_{agg}^{no} \quad \textit{and} \quad M\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = 0$$

with probability 1, and

$$Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) = y_{agg}^{no} - \frac{\psi\epsilon n}{\delta} \quad \textit{and} \quad M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) = -\frac{\psi n}{\delta}$$

with probability 1.

Proposition 2.13 nests the setting of Proposition 2.12 if we set $\delta = \sigma$. When $\mathbf{1}^T\mathbf{\Sigma} = \delta\mathbf{1}^T$, it turns out that $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) = -\frac{1}{N\delta}\mathbf{1}$, and every agent has the same weight. As a result, regardless of which subset of agents receives a positive monetary transfer or stimulus, the aggregate action remains the same. The distribution of aggregate actions and the distribution of economic multipliers are both degenerate. Agents' interaction structure needs to deviate from this null setting in order to obtain non-degenerate distributions. For null interaction structures, $M\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = 0$, so any deviation from this class of network topologies leads to the emergence of negative economic multipliers in a setting with transfers.

For the remainder of this section, we provide a set of results that allows us to rank networks. A higher-ranked network generates distributions of aggregate actions and/or distributions of multipliers that first-order stochastically dominate those generated by a lower-ranked network. For every possible configuration of transfers or stimulus, we find that the higher-ranked network generates a higher level of aggregate actions and/or a larger economic multiplier than a lower-ranked network. Transfers and stimulus are therefore relatively more effective for the higher-ranked network. Our first result focuses on aggregate actions. It ranks networks according to their corresponding distributions of aggregate actions:

Proposition 2.14 *Provided that $\beta > \lambda r(\mathbf{G})$, $\beta' > \lambda' r(\mathbf{G}')$, and $\gamma' = 0$, when $\mathbf{\Sigma}' > \mathbf{\Sigma}$, $Y_{agg}\left(-(\mathbf{\Sigma}')^{-1}, N, n, 0\right) \succeq Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right)$ and $Y_{agg}\left(-(\mathbf{\Sigma}')^{-1}, N, n, 1\right) \succeq Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ for all $n \in \{1, \ldots, N-1\}$.*

Proposition 2.14 requires that $\gamma' = 0$, which means that there can be no strategic substitutes in the $\mathbf{\Sigma}'$ environment. The $\mathbf{\Sigma}$ environment can admit both strategic complements and strategic substitutes. In general, when $\mathbf{\Sigma}' > \mathbf{\Sigma}$ and $\gamma' = 0$, the distribution of possible aggregate actions in the $\mathbf{\Sigma}'$ environment first-order stochastically dominates the distribution of possible aggregate actions in the $\mathbf{\Sigma}$ environment. This result separately holds in settings with transfers and in settings with stimulus. To prove this result, we demonstrate that the aggregate action in the $\mathbf{\Sigma}'$ environment exceeds the aggregate action in the $\mathbf{\Sigma}$ environment for any wealth vector $\boldsymbol{\omega} + \boldsymbol{\rho}$; this wealth vector can represent the wealth of agents in settings with transfers and it can represent the wealth of agents in settings with stimulus.

Our second result focuses on economic multipliers. In settings with transfers, we are unable to rank networks so that the distribution of multipliers for the higher-ranked network first-order stochastically dominates the distribution of multipliers for the lower-ranked network. This is because $EM\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = 0$; regardless of the underlying network the mean multiplier is always the same. However, in settings with stimulus, we can rank networks according to their corresponding distributions of economic multipliers:

Proposition 2.15 *Provided that $\beta > \lambda r(\mathbf{G})$, $\beta' > \lambda' r(\mathbf{G}')$, and $\gamma' = 0$, $M\left(-(\mathbf{\Sigma}')^{-1}, N, n, 1\right) \succeq M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ for all $n \in \{1, \ldots, N\}$ when $\mathbf{\Sigma}' > \mathbf{\Sigma}$ for symmetric $\mathbf{\Sigma}, \mathbf{\Sigma}'$.*

The $\mathbf{\Sigma}'$ environment only admits strategic complements, while the $\mathbf{\Sigma}$ environment can admit both strategic complements and strategic substitutes. Under these assumptions, $\mathbf{d}_w^-\left(-(\mathbf{\Sigma}')^{-1}\right) > \mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)$, which makes the distribution of multipliers in the $\mathbf{\Sigma}'$ environment first-order stochastically dominate the distribution of multipliers in the $\mathbf{\Sigma}$ environment. Given that n agents each receive $\epsilon > 0$ units of stimulus, the effect on the aggregate action is relatively more positive for the higher-ranked network.

2.4 Networked Environments with Coordination and Anti-Coordination

We proceed to our second environment with network-based interaction. The environment that we focus on in this section is a dynamic one in which agents engage in a mixture of coordinating and anti-coordinating behavior with other agents in the population. We are interested in the aggregate action for the population and its dynamic evolution.

We have a population of N agents. Each agent chooses an action that somehow depends on other agents' past behaviors. In choosing this action, each agent essentially segments the population into two groups: (1) a group with whom the agent seeks to choose a coordinating action and (2) a group with whom the agent seeks to choose an anti-coordinating action. Each agent moreover decides how much weight to accord to every other agent in the population. The $N \times N$

matrix $\mathbf{T}$ captures each agent's desire for coordination or anti-coordination; it is a matrix of linkage types. Similar to Eger (2016a), there are two types of linkages: $[\mathbf{T}]_{ij} \in \{\mathcal{F}, \mathcal{D}\}\ \forall i, j \in \{1, \ldots, N\}$. $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ is the *follow* linkage, while $\mathcal{D} : \mathbb{R} \to \mathbb{R}$ is the *deviation* linkage. Given the past action of an agent j, function $\mathcal{F}$ or $\mathcal{D}$ transforms that past action into the present desired responding action for agent i. When $[\mathbf{T}]_{ij} = \mathcal{F}$, agent i seeks to follow the past action of agent j, so we have myopic coordination, while when $[\mathbf{T}]_{ij} = \mathcal{D}$, agent i seeks to deviate from the past action of agent j, so we have myopic anti-coordination. Later on, we will be introducing a matrix $\mathbf{O}$, with $[\mathbf{O}]_{ij} = 1$ if $[\mathbf{T}]_{ij} = \mathcal{F}$ and $[\mathbf{O}]_{ij} = -1$ if $[\mathbf{T}]_{ij} = \mathcal{D}$. Meanwhile, the $N \times N$ row-stochastic matrix $\bar{\mathbf{A}}$ captures the weight that agents assign to other agents. The ij^{th} element of $\bar{\mathbf{A}}$ represents the non-negative weight that agent i allocates to agent j. For every agent $i \in \{1, \ldots, N\}$, the sum of the weights that each agent i accords to every other agent j sums to 1, that is, $[\bar{\mathbf{A}}]_{i*} \mathbf{1} = 1$, and unless otherwise specified, $[\bar{\mathbf{A}}]_{ii} = 0$. In our environment with coordinating and anti-coordinating behavior, agents are therefore organized on the network $\mathcal{G}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)$ with corresponding weighted adjacency matrix $\bar{\mathbf{A}} \circ \mathbf{O}$.

Every period q, agent i chooses an action, $y_{i,q}$, that maximizes his period-q utility:

$$\max_{y_{i,q}} u_{i,q} = \max_{y_{i,q}} -\sum_{j=1}^{N} [\bar{\mathbf{A}}]_{ij} \left(y_{i,q} - [\mathbf{T}]_{ij} \left(y_{j,q-1}\right)\right)^2.$$

We define $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ to be an identity function; $\mathcal{F}\left(y_{j,q-1}\right) = y_{j,q-1}$, so when $[\mathbf{T}]_{ij} = \mathcal{F}$, agent i seeks to choose an action that follows agent j's past action. We define $\mathcal{D} : \mathbb{R} \to \mathbb{R}$ as $\mathcal{D}\left(y_{j,q-1}\right) = y - \left(y_{j,q-1} - y\right)$. When $[\mathbf{T}]_{ij} = \mathcal{D}$, agent i seeks to choose an action that deviates from the past action of agent j. In particular, agent i wishes to choose an action that deviates in a direction opposite to the previous action of agent j; for example, if the past action of agent j is less than a benchmark action y, then agent i seeks to choose an action that is greater than y, and vice versa. We assume that desired deviating behavior takes the same form across all agents. In this setting, coordination and anti-coordination occur on past actions, so agents' behavior is myopic.

There are different ways that we can think about this theoretical environment and how it maps to realistic settings. Here, I focus on one particular mapping. We can imagine that there is a population of N agents who choose an action every period. The magnitude of the action that each agent chooses depends on the prior actions of other agents. In particular, each agent has a set of role models and anti-role models. Matrix element $[\mathbf{O}]_{ij} = 1$ if agent j is a role model for agent i, and matrix element $[\mathbf{O}]_{ij} = -1$ if agent j is instead an anti-role model for agent i. Matrix $\bar{\mathbf{A}}$ then captures the weight that each agent accords to his or her role models and anti-role models. In a setting with role models and anti-role models, agents respectively engage in myopic coordination and anti-coordination. In general, agents do not communicate with their role models and anti-role models. Rather, they observe the past actions of these agents, and they then seek to choose an action that imitates the past actions of their role models and deviates from the past actions of their anti-role models. A setting with role models and anti-role models is therefore a natural setting for myopic coordination and anti-coordination. Now, there is a wide range of possible actions that these agents can take. Let's assume that agents are engaging in prosocial behavior, such as volunteering or providing a public good. These agents are selecting the amount of time that they engage in this activity, with the amount of time dependent on the past actions of other agents. The outside observer to this system is interested in the aggregate action, which is the total amount of time spent on the activity.

We next return to our model and solve the optimization problem for each agent in this environment, that is, the optimal choice of a period-specific action:

Proposition 2.16 *For each agent* $i \in \{1, \ldots, N\}$, $y^*_{i,q} = \sum_{j=1}^{N} [\bar{\mathbf{A}}]_{ij} [\mathbf{T}]_{ij} \left(y_{j,q-1}\right)$, *and therefore* $\mathbf{y}^*_q = \left(\bar{\mathbf{A}} \circ \mathbf{T}\right)^q \mathbf{y}_0$.

Depending on the structure of $[\mathbf{T}]_{i*}$, agent i engages in a mixture of coordinating and anti-coordinating behavior. Agent i chooses an action that is a weighted sum of other agents' past actions, for those agents that agent i seeks to follow, and a weighted sum of desired deviating actions, for those agents from whom agent i seeks to deviate. Agents coordinate and anti-coordinate on past actions. They have myopic best-response functions.

We assume that all agents prior to time period zero choose action y. This action is an optimal action; given that every agent chooses y in period $q-1$, every agent will continue to choose y in period q:

$$y^*_{i,q} = \sum_{j=1}^{N} [\bar{\mathbf{A}}]_{ij} [\mathbf{T}]_{ij} \left(y_{j,q-1}\right)$$

$$y^*_{i,q} = \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij}\, y + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij} \left(2y - y\right) = y.$$

When time period 0 arrives, an outside entity adjusts agents' actions: $\mathbf{y}^*_0 = y\mathbf{1} + \boldsymbol{\rho}$; the $\boldsymbol{\rho}$ vector captures that adjustment. In settings with transfers, the outside entity increases the actions of n agents by $\epsilon > 0$ units, and decreases the actions of the remaining $N-n$ agents by $\frac{n\epsilon}{N-n}$ units. Therefore, if the i^{th} agent receives a positive transfer, $[\boldsymbol{\rho}]_i = \epsilon$, while if the i^{th} agent receives a negative transfer, $[\boldsymbol{\rho}]_i = -\frac{n\epsilon}{N-n}$. In settings with stimulus, the outside entity only increases the actions of n agents by $\epsilon > 0$ units; it does not adjust the period-0 actions of the remaining $N-n$ agents. Therefore, if the i^{th} agent receives positive stimulus, $[\boldsymbol{\rho}]_i = \epsilon$, while if the i^{th} agent does not receive positive stimulus, then $[\boldsymbol{\rho}]_i = 0$. Configuration vector

$\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ determines which agents receive that positive period-0 shock in both settings with transfers and settings with stimulus. The next result allows us to analytically trace the population vector of optimal agent actions for all periods q given the period-0 shock $\boldsymbol{\rho}$:

Proposition 2.17 *With $\mathbf{y}_q^* = y\mathbf{1}$ for $q < 0$, the population vector of agent actions is $\mathbf{y}_q^* = y\mathbf{1} + (\bar{\mathbf{A}} \circ \mathbf{O})^q \boldsymbol{\rho}$ for all $q \in \mathbb{Z}_+$.*

We can now compute the aggregate action in period q:

$$y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, \mathbf{b}, N, n, \ell\right) = \mathbf{1}^T \mathbf{y}_q^* = y_{agg}^{no} + N\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T \boldsymbol{\rho}$$

for $\ell \in \{0, 1\}$.[2] Quantity $y_{agg}^{no} = Ny$ is the aggregate action absent transfers or stimulus, that is, when $\boldsymbol{\rho} = \mathbf{0}$. $\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$ is the vector of average weighted in-degrees for graph $\mathcal{G}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$. We set $\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right) = \frac{1}{N}\left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]^T \mathbf{1}$. The sum of agents' weights in period q is k_q: $\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T \mathbf{1} = k_q$.

In this particular environment, the aggregate action depends on the structure of agents' interaction network. We essentially have two relevant networks: (1) the original agent interaction structure, $\mathcal{G}(\mathbf{Z}') = \mathcal{G}(\bar{\mathbf{A}} \circ \mathbf{O})$, that captures agents' myopic coordinating and anti-coordinating behavior, and (2) the network, $\mathcal{G}(\mathbf{Z}) = \mathcal{G}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$, that determines each agent's effective weight in the population. Each agent's weight identifies how much of an effect targeted stimulus or a targeted transfer towards that particular agent has on the aggregate action. We have this complicated mixture of coordinating and anti-coordinating behavior among agents that increases in complexity as time evolves. The period-specific vector of agent weights, $\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$, summarizes the net amount of coordination

[2] The notation in this section deviates slightly from the notation introduced in Chapter 1. To be consistent with the other sections in the present chapter, we use $\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$ instead of $\mathbf{d}_w^{-(q)}(\bar{\mathbf{A}} \circ \mathbf{O})$, and later on, we use $\widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)$ instead of $\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}} \circ \mathbf{O}, N, n)$.

or anti-coordination that the entire population of agents undertakes given the action of every agent.

For a particular configuration of transfers or stimulus, we can compute the aggregate action, the corresponding dynamic multiplier (i.e., the period-specific economic multiplier), and the impulse response. First, let's suppose that agents $1, \ldots, n$ receive a positive transfer. The aggregate action in period q is then:

$$\begin{aligned}
y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 0\right) = \\
& y_{agg}^{no} + N\epsilon\left(\left[\mathbf{d}_{w}^{-}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right)\right]_{1} + \cdots + \left[\mathbf{d}_{w}^{-}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right)\right]_{n}\right) \\
& \quad - N\frac{n\epsilon}{N-n}\left(\left[\mathbf{d}_{w}^{-}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right)\right]_{n+1} + \cdots + \left[\mathbf{d}_{w}^{-}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right)\right]_{N}\right),
\end{aligned}$$

The period-specific dynamic multiplier is

$$m_{q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 0\right) = \frac{dy_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 0\right)}{d\epsilon},$$

and the impulse response function is

$$irf_{q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 0\right) = y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 0\right) - y_{agg}^{no}.$$

Next, let's suppose that agents $1, \ldots, n$ instead receive positive stimulus. The aggregate action in period q is then:

$$\begin{aligned}
y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 1\right) = y_{agg}^{no} + N\epsilon\left(\left[\mathbf{d}_{w}^{-}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right)\right]_{1} + \cdots\right. \\
\left. + \left[\mathbf{d}_{w}^{-}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right)\right]_{n}\right).
\end{aligned}$$

The period-specific dynamic multiplier is

$$m_{q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 1\right) = \frac{dy_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^{q}, \mathbf{b}, N, n, 1\right)}{d\epsilon},$$

and the impulse response function is

$$irf_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,\mathbf{b},N,n,1\right)=y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,\mathbf{b},N,n,1\right)-y_{agg}^{no}.$$

Depending on which subset of agents receives a positive transfer or stimulus, we can have wide variation in the aggregate action, the economic multiplier, and the impulse response for each period. In the next proposition, we define the random variables that allow us to construct these distributions of possible values for the aggregate action, the economic multiplier, and the impulse response for every period q:

Proposition 2.18 *In a setting with transfers, the aggregate action, the dynamic multiplier, and the impulse response are:*

$$Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right)=y_{agg}^{no}+\frac{N^2\epsilon}{N-n}\left(\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right)-\frac{k_q n}{N}\right),$$
$$M_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right)=\frac{N^2}{N-n}\left(\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right)-\frac{k_q n}{N}\right),\ \text{and}$$
$$IRF_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right)=\frac{N^2\epsilon}{N-n}\left(\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right)-\frac{k_q n}{N}\right).$$

In a setting with stimulus, the aggregate action, the dynamic multiplier, and the impulse response are:

$$Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right)=y_{agg}^{ns}+N\epsilon\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right),$$
$$M_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right)=N\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right),\ \text{and}$$
$$IRF_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right)=N\epsilon\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right).$$

The random variables for the impulse response function, $IRF_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right)$ and $IRF_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right)$, are defined in a similar manner to the other random variables in Proposition 2.18.[3]

[3]Provided that each period-0 configuration of transfers or stimulus is equally likely, the CDF for

The expressions for $Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},N,n,0\right)$, $M_{q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},N,n,0\right)$, $Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},N,n,1\right)$, and $M_{q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},N,n,1\right)$ in Proposition 2.18 directly map to Equations 2.2-2.5 in Section 2.2, where we set $\mathbf{Z}=(\bar{\mathbf{A}}\circ\mathbf{O})^{q}$ and $\gamma_1=1$. Accordingly, we can map all of the results from Section 2.2 to the present theoretical environment. We can characterize the period-specific distributions of aggregate actions, dynamic economic multipliers, and impulse responses. We can solve for the moments of these distributions in closed form. We can determine the lowest possible aggregate action and the highest possible aggregate action for a particular level of transfers or stimulus, and we can demonstrate how those values depend on the topology of agents' interaction network. Based on the asymptotic expansions that approximate these distributions, we can determine the extent to which agents' interaction topology leads to skewness and/or heavy-tailedness in the distributions of possible aggregate actions, dynamic economic multipliers, and impulse responses.

We can also importantly compute both the probability that the dynamic multiplier is negative and the probability that the aggregate action drops below its no-intervention level, y_{agg}^{no}, for finite q and in the limiting case as $q\rightarrow\infty$. In this environment with myopic coordination and anti-coordination, there are different reasons why negative multipliers can arise. We now illustrate a couple of these pathways. Let's first consider a setting with transfers. The aggregate action can decline if an agent receives a positive shock and other agents seek to anti-coordinate with that particular agent. The aggregate action can separately decline if an agent receives a negative shock and other agents seek to coordinate

---

$IRF_{q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},N,n,\ell\right)$ is

$$G_{IRF_{q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},N,n,\ell\right)}(t)=\frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)}\mathbb{1}_{irf_{q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^{q},\mathbf{b},N,n,\ell\right)\leq t}\quad\text{for }\ell\in\{0,1\}.$$

with that particular agent. Let's next consider a setting with stimulus. Every agent is receiving a non-negative shock at period zero. If we only have coordinating behavior, then there is zero probability that the multiplier can be negative. However, if we have anti-coordinating behavior, the aggregate action can decline if the desire to anti-coordinate with an agent receiving stimulus is sufficiently strong. For the mechanisms just described, agents are adjusting their actions upward or downward based on the shocks that their immediate neighbors receive. However, as time evolves, even though agents are continuing to myopically coordinate and anti-coordinate with their network neighbors, they are indirectly coordinating and anti-coordinating with agents whose distance on the network exceeds 1.

We would like to characterize the period-specific aggregate action, dynamic multiplier, and impulse response in two different null settings. For the first null setting, there is no underlying network:

Proposition 2.19 *In the absence of network-based interaction,* $Y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 0\right) = y_{agg}^{no}$*,* $M_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 0\right) = 0$*, and* $IRF_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 0\right) = 0$ *with probability 1 for all* $q \in \mathbb{Z}_+$*, and* $Y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = y_{agg}^{no} + n\epsilon$*,* $M_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = n$*, and* $IRF_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = n\epsilon$ *with probability 1 for all* $q \in \mathbb{Z}_+$*.*

When there is no network, $\bar{\mathbf{A}} \circ \mathbf{O} = \mathbf{I}$, so agents coordinate on their past actions. In settings with transfers, the aggregate action equals its no-intervention level for all possible initial configurations and all periods $q \in \mathbb{Z}_+$. In settings with stimulus, the aggregate action exceeds its no-intervention level by the total amount of stimulus, $n\epsilon$, and the aggregate action maintains this value for all possible initial configurations of stimulus and all periods $q \in \mathbb{Z}_+$. Non-trivial network-based interaction enables us to have a non-degenerate distribution for the aggregate action; in such an environment, the aggregate action can deviate in either direction away from its no-intervention level.

For the second null setting, the aggregate action is invariant to the particular configuration of transfers or stimulus. We identify the necessary restrictions on $(\bar{\mathbf{A}} \circ \mathbf{O})^q$ that make the period-q aggregate action invariant to configuration, and we then solve for the resulting values of the aggregate action, the corresponding economic multiplier, and the impulse response:

Proposition 2.20 *If $\mathbf{1}^T (\bar{\mathbf{A}} \circ \mathbf{O})^q = k_q \mathbf{1}^T$, the period-$q$ aggregate action, dynamic multiplier, and impulse response are invariant to configuration: $Y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 0\right) = y_{agg}^{no}$, $M_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 0\right) = 0$, and $IRF_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 0\right) = 0$ with probability 1, and $Y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = y_{agg}^{no} + k_q n \epsilon$, $M_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = k_q n$, and $IRF_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = k_q n \epsilon$ with probability 1.*

Proposition 2.20 nests the setting of Proposition 2.19 if we set $k_q = 1$. When $\bar{\mathbf{A}}$ is doubly stochastic and $\bar{\mathbf{A}} \circ \mathbf{O} = \bar{\mathbf{A}}$, that is, the environment only has coordination, $\mathbf{1}^T (\bar{\mathbf{A}} \circ \mathbf{O})^q = \mathbf{1}^T$ for all $q \in \mathbb{Z}_+$, which makes the aggregate action, economic multiplier, and impulse response invariant to configuration for all time periods $q \in \mathbb{Z}_+$. For a transfer or stimulus to generate a negative economic multiplier, we must deviate from those network topologies characterized in Proposition 2.20, for which the economic multiplier is invariant to configuration.

The next result allows us to rank networks so that the distributions of aggregate actions, dynamic multipliers, and impulse responses for the higher-ranked network first-order stochastically dominate the distributions of aggregate actions, dynamic multipliers, and impulse responses for the lower-ranked network:

Proposition 2.21 *If $(\bar{\mathbf{A}}' \circ \mathbf{O}')^q \geq (\bar{\mathbf{A}} \circ \mathbf{O})^q \mathbf{P}$ for some permutation matrix $\mathbf{P}$, then $Y_{agg,q}\left((\bar{\mathbf{A}}' \circ \mathbf{O}')^q, N, n, 1\right) \succeq Y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right)$, $M_q\left((\bar{\mathbf{A}}' \circ \mathbf{O}')^q, N, n, 1\right) \succeq M_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right)$, and $IRF_q\left((\bar{\mathbf{A}}' \circ \mathbf{O}')^q, N, n, 1\right) \succeq IRF_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right)$ for all $n \in \{1, \ldots, N-1\}$.*

We can rank network topologies in settings with stimulus. Proposition 2.21 nests the case in which a linkage changes from negative to positive; in such an environment, $\bar{\mathbf{A}}' = \bar{\mathbf{A}}$, but $\mathbf{O}' > \mathbf{O}$. Once we move from anti-coordination to coordination for one pair of agents, the distribution of aggregate actions first-order stochastically dominates the original distribution. We are not able to rank network topologies in settings with transfers because the distribution of aggregate actions always has a mean equal to its no-intervention level, y_{agg}^{no}, for every feasible network structure.

Now that we have ranked networks, we proceed to determine the maximum and minimum possible dynamic multipliers among all networks. First, we must bound the allowable values for k_q:

Lemma 2.7 *For all* $q \geq 1$, $k_q \in [-1, 1]$.

$k_q = 1$ is attainable when $\bar{\mathbf{A}} \circ \mathbf{O} = \bar{\mathbf{A}}$, and $k_q = -1$ is attainable when $\bar{\mathbf{A}} \circ \mathbf{O} = -\bar{\mathbf{A}}$. We must also introduce two classes of graphs:

Definition 2.3 *Graph* $\mathcal{G}(\mathbf{Z})$ *is a* positive star graph *and a* negative star graph *when the weighted adjacency matrices are respectively:*

$$\mathbf{Z} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \mathbf{P} \quad \textit{and} \quad \mathbf{Z} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -1 \\ 0 & 0 & \cdots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix} \mathbf{P}$$

for some $N \times N$ *permutation matrix* $\mathbf{P}$*.*

In the next proposition, we identify the extremal multipliers in settings with stimulus. These extremal values hold for all levels n, that is, for all possible levels of stimulus. We also identify the network topologies that generate these extremal multipliers, temporarily allowing for the existence of self-loops:

Proposition 2.22 *For every $n \in \{1, \ldots, N-1\}$ and all feasible matrices $(\bar{\mathbf{A}} \circ \mathbf{O})^q$,*

$$\max_{(\bar{\mathbf{A}} \circ \mathbf{O})^q} \left[\max \operatorname{supp} M_q \left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right)\right] = N \text{ and}$$

$$\min_{(\bar{\mathbf{A}} \circ \mathbf{O})^q} \left[\min \operatorname{supp} M_q \left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right)\right] = -N.$$

These maximum and minimum values are attainable when graph $\mathcal{G}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$ is respectively a positive star graph and a negative star graph.

Among all networks, in settings with stimulus, the maximum possible dynamic multiplier is N, and among all networks, in settings with stimulus, the minimum possible dynamic multiplier is $-N$. These maximum and minimum values are attainable for every level of stimulus, that is, for every integer $n \in \{1, \ldots, N-1\}$. If $\mathcal{G}(\bar{\mathbf{A}} \circ \mathbf{O})$ is a positive star graph, then the maximum possible dynamic multiplier is attainable for all periods $q \geq 1$ and for all levels $n \in \{1, \ldots, N-1\}$ of initial stimulus due to the idempotence of $\bar{\mathbf{A}} \circ \mathbf{O}$.

We conclude this section by noting that our environment with myopic coordination and anti-coordination is a dynamic one. Provided that $\bar{\mathbf{A}} \circ \mathbf{O}$ is semi-convergent, $\lim_{q \to \infty} (\bar{\mathbf{A}} \circ \mathbf{O})^q$ exists and we can characterize limiting distributions for the aggregate action, dynamic multiplier, and impulse response in settings with transfers and settings with stimulus as $q \to \infty$. From our random variables in Proposition 2.18, we see that these limiting distributions all depend on the following object: $\lim_{q \to \infty} \widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)$. In settings with transfers, the limiting distributions depend on the quantity $\lim_{q \to \infty} k_q$ as well. To compute $\lim_{q \to \infty} \widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)$, we must determine the limiting vector $\lim_{q \to \infty} \mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$.

The remaining theoretical results allow us to solve for $\lim_{q \to \infty} \mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$ after making specific assumptions. Let us first consider the case in which we only have coordination, so that $\bar{\mathbf{A}} \circ \mathbf{O} = \bar{\mathbf{A}}$:

Proposition 2.23 *If* $\bar{\mathbf{A}}$ *is primitive, then* $\lim_{q\to\infty}\mathbf{d}_w^-\left(\bar{\mathbf{A}}^q\right) = \mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)$*, where the pair* $\left(\mathbf{w}_\infty^T, 1\right)$ *is the unique dominant left eigenpair of* $\bar{\mathbf{A}}$*,* $\mathbf{w}_\infty^T\bar{\mathbf{A}} = \mathbf{w}_\infty^T$*, and* $\mathbf{w}_\infty^T\mathbf{1} = 1$*.*

The limiting vector of agent weights is computed by solving for the left eigenvector of the matrix $\bar{\mathbf{A}}$ that pairs with the unit eigenvalue. Provided that $\bar{\mathbf{A}}$ is primitive, when $\bar{\mathbf{A}} \circ \mathbf{O} = \bar{\mathbf{A}}$, all agents converge to the same limiting action because $\lim_{q\to\infty}\bar{\mathbf{A}}^q = \mathbf{1}\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T$. From the row-stochasticity of $\bar{\mathbf{A}}$, we are also able to determine that $\lim_{q\to\infty} k_q = 1$.

Given some additional assumptions, in the coordinating environment we can compute the probability of a negative limiting dynamic multiplier in terms of the network's primitives. Let us assume that $\left[\bar{\mathbf{A}}\right]_{ij} > 0$ if and only if $\left[\bar{\mathbf{A}}\right]_{ji} > 0$ and agents assign an equal weight to all out-neighbors. Since all linkages in the network are accordingly reciprocal, we can compute a vector of degrees, $\mathbf{d}\left(\bar{\mathbf{A}}\right)$. The degree for agent i is equal to the number of non-zero elements in the i^{th} row of $\bar{\mathbf{A}}$: $\left[\bar{\mathbf{A}}\right]_{i*}$. Given the vector of degrees, we define a random variable $D\left(\bar{\mathbf{A}}\right)$ whose realization is the degree for agent i: $\left[\mathbf{d}\left(\bar{\mathbf{A}}\right)\right]_i$.

Proposition 2.24 *Suppose that* $\bar{\mathbf{A}}$ *is primitive,* $\left[\bar{\mathbf{A}}\right]_{ij} > 0$ *if and only if* $\left[\bar{\mathbf{A}}\right]_{ji} > 0$*, and all non-zero elements within every row of* $\bar{\mathbf{A}}$ *have the same value. When* $n = 1$*,*
$\Pr\left[\lim_{q\to\infty} Y_{agg,q}\left(\bar{\mathbf{A}}^q, N, n, 0\right) < y_{agg}^{no}\right] = \Pr\left[\lim_{q\to\infty} M_q\left(\bar{\mathbf{A}}^q, N, n, 0\right) < 0\right] =$
$\Pr\left[D\left(\bar{\mathbf{A}}\right) < ED\left(\bar{\mathbf{A}}\right)\right]$.

When one agent receives a positive transfer at time period 0, the probability that the limiting dynamic multiplier ends up being negative is equal to the probability that the degree in the network is less than the expected degree.[4] Meanwhile, in a

[4] We can construct propositions similar to Proposition 2.24 for networks $\mathcal{G}\left(\bar{\mathbf{A}}\right)$ that have other sets of features as well. Refer to Section 1.4 of Chapter 1 for different closed-form formulations of $\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)$.

setting with stimulus, since agents are all coordinating, the probability of a negative limiting dynamic multiplier is always zero.

Lastly, we consider the case in which we have a mixture of coordinating and anti-coordinating behavior, so that $\bar{\mathbf{A}} \circ \mathbf{O} \neq \bar{\mathbf{A}}$. $\lim_{q\to\infty} \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right)$ exists when $\bar{\mathbf{A}} \circ \mathbf{O}$ is semi-convergent. Under certain assumptions, we can explicitly solve for $\lim_{q\to\infty} \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right)$. Let graph $\mathcal{G} (\bar{\mathbf{A}} \circ \mathbf{O})$ have edge weight $e_{i,j} = [\bar{\mathbf{A}} \circ \mathbf{O}]_{ij}$. We characterize $\lim_{q\to\infty} \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right)$ after providing definitions for structural balance and absolute row-stochasticity:

Definition 2.4 *Graph $\mathcal{G} (\bar{\mathbf{A}} \circ \mathbf{O}) = (\mathcal{V} (\bar{\mathbf{A}} \circ \mathbf{O}), \mathcal{E} (\bar{\mathbf{A}} \circ \mathbf{O}))$ is* structurally balanced *if there exists a partition of nodes $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, with $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, such that: (1) if nodes $i, j \in \mathcal{V}_\ell$ and $(i, j, e_{i,j}) \in \mathcal{E}$ for $\ell \in \{1, 2\}$, then $\operatorname{sgn}(e_{i,j}) > 0$, and (2) if $i \in \mathcal{V}_\ell$ and $j \in \mathcal{V}_{-\ell}$ with $(i, j, e_{i,j}) \in \mathcal{E}$ for $\ell \in \{1, 2\}$, then $\operatorname{sgn}(e_{i,j}) < 0$.*[5]

Definition 2.5 *$\bar{\mathbf{A}} \circ \mathbf{O}$ is* absolutely row-stochastic *if $|\bar{\mathbf{A}} \circ \mathbf{O}|$, the element-wise absolute value of matrix $\bar{\mathbf{A}} \circ \mathbf{O}$, is row-stochastic.*

We now solve for $\lim_{q\to\infty} \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right)$:

Proposition 2.25 *For $\bar{\mathbf{A}} \circ \mathbf{O}$ absolutely row-stochastic, $|\bar{\mathbf{A}} \circ \mathbf{O}| = \bar{\mathbf{A}}$ primitive, and $\mathcal{G} (\bar{\mathbf{A}} \circ \mathbf{O})$ structurally balanced, $\lim_{q\to\infty} \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right) = \lim_{q\to\infty} \frac{1}{N} \left[(\bar{\mathbf{A}} \circ \mathbf{O})^q \right]^T \mathbf{1}$ exists, with $\lim_{q\to\infty} (\bar{\mathbf{A}} \circ \mathbf{O})^q = \left(\mathbf{1} \left[\mathbf{w}_\infty (\bar{\mathbf{A}}) \right]^T \right) \circ \mathbf{O}$ and $\left[\mathbf{w}_\infty (\bar{\mathbf{A}}) \right]^T \bar{\mathbf{A}} = \left[\mathbf{w}_\infty (\bar{\mathbf{A}}) \right]^T$.*

If $|\mathcal{V}_1| < |\mathcal{V}_2|$, then $\lim_{q\to\infty} \left[\mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right) \right]_i < 0$ if $i \in \mathcal{V}_1$ and $\lim_{q\to\infty} \left[\mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right) \right]_i > 0$ if $i \in \mathcal{V}_2$. Since $\lim_{q\to\infty} \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q \right) \neq \lim_{q\to\infty} \frac{k_q}{N} \mathbf{1}$, in settings with transfers, we have a positive probability of a negative limiting dynamic multiplier for all levels of transfers, that is, for all $n \in \{1, \ldots, N-1\}$.

[5] Eger (2016b) refers to this property as +-opposition bipartiteness.

2.5 Networked Environments with Production

In this third and final setting with network-based interaction among agents, we consider a production network. We have a population of firms; every firm represents a different sector. Production by each firm potentially requires both labor and intermediate goods obtained from other firms. Linkages in the production network therefore capture the flow of intermediate goods between firms. Specifically, the directed edges of the network designate downstreamness in production; for each firm, these edges point towards the firm's customers.

Now, given the production network, we are interested in the nominal value of aggregate output, or GDP. In particular, we are interested in the distribution of possible levels of GDP that result when a random subset of sectors receives a positive demand shock. Sectors differ in their importance within the production network, so depending on which group of sectors receives a positive demand shock, we can potentially have strong variation in GDP. We are also interested in the distribution of possible economic multipliers. We would like to know the change in GDP that results when a certain subset of sectors receives a positive demand shock of a particular magnitude. Depending on the structure of the production network, we can imagine that the multiplier on GDP varies with the group of sectors actually receiving the positive demand shock. For certain groups of sectors, the boost in GDP is larger than that for other groups.

We proceed to describe this section's theoretical environment. There are N sectors with one good associated with each sector, so there are N total goods. Our representative consumer has Cobb-Douglas utility over the consumption bundle $\mathbf{c} = \begin{pmatrix} c_1 & \cdots & c_N \end{pmatrix}^T$:

$$U(\mathbf{c}) = \prod_{i=1}^{N} c_i^{\eta_i}.$$

We set $\sum_{i=1}^{N} \eta_i = 1$. Each good is produced by a competitive sector; the good is either consumed or used in the production of other goods. The production technology for the representative firm in each sector takes a Cobb-Douglas form:

$$x_i = A_i^{\alpha_i} \ell_i^{\alpha_i} \left(\prod_{j=1}^{N} x_{ji}^{[\boldsymbol{\Lambda}]_{ji}} \right)^{\beta_i}.$$

A_i is a sector-specific productivity parameter; productivity is labor-augmenting. ℓ_i denotes the amount of labor used in the production of good i. The representative consumer has no disutility from labor, so labor is inelastically supplied. x_{ji} denotes the quantity of the sector-j good required in the production of the sector-i good. Exponent $[\boldsymbol{\Lambda}]_{ji}$ captures how intensely the sector-j good is used in the production of the sector-i good; $[\boldsymbol{\Lambda}]_{ji}$ represents the share of good j in total intermediate input use by sector i. We assume that the production technology for each sector is constant returns to scale, that is, $\alpha_i + \beta_i = 1 \ \forall i \in \{1, \ldots, N\}$. Parameters α_i and β_i can differ across sectors; sectors vary in the intensities with which they use labor and intermediate inputs. We also assume that $\boldsymbol{\Lambda}$ is column-stochastic; this ensures that the overall sectoral production function is constant returns to scale. The production network in this environment is $\mathbf{Z}' = \mathcal{G}(\boldsymbol{\Lambda})$.

In this economy, each firm $i \in \{1, \ldots, N\}$ maximizes its profit π_i: $\pi_i = p_i x_i - \sum_{j=1}^{N} p_j x_{ji} - w\ell_i$, where p_i is the price of good i and w is the wage rate. The representative consumer maximizes its utility subject to a resource constraint: $\sum_{i=1}^{N} p_i c_i = w \sum_{i=1}^{N} \ell_i + \sum_{i=1}^{N} \pi_i$. Labor market clearing requires that the supply of labor equals the total demand for labor: $1 = \sum_{i=1}^{N} \ell_i$, where we set the supply of labor equal to 1. Goods market clearing requires that, for each sector $i \in \{1, \ldots, N\}$, the supply of the good equals the demand for that good: $x_i = c_i + \sum_{j=1}^{N} x_{ij}$. We define the matrix $\mathbf{X}$ with elements x_{ij}, the quantity of the sector-i good used in the production of the good from sector j.

A competitive equilibrium in this economy can be characterized as follows:

Definition 2.6 *A* competitive equilibrium *is a collection of quantities,* $\mathbf{c}^*$, $\mathbf{x}^*$, $\mathbf{X}^*$, *and* $\boldsymbol{\ell}^*$, *and a collection of prices,* $\mathbf{p}^*$ *and* w^*, *such that:*

1. *The representative consumer maximizes utility subject to a budget constraint:*

$$\max_{c_1,\ldots,c_N} \prod_{i=1}^{N} c_i^{\eta_i} \quad s.t. \quad \sum_{i=1}^{N} p_i c_i = w \sum_{i=1}^{N} \ell_i + \sum_{i=1}^{N} \pi_i.$$

2. *Each firm* $i \in \{1,\ldots,N\}$ *maximizes profit given its production technology:*

$$\max_{x_{1i},\ldots,x_{Ni},\ell_i} \; p_i x_i - \sum_{j=1}^{N} p_j x_{ji} - w\ell_i$$

$$s.t. \quad x_i = A_i^{\alpha_i} \ell_i^{\alpha_i} \left(\prod_{j=1}^{N} x_{ji}^{[\boldsymbol{\Lambda}]_{ji}} \right)^{\beta_i} \; \forall i \in \{1,\ldots,N\}.$$

3. *The goods markets clear for all sectors* $i \in \{1,\ldots,N\}$ *and the labor market clears:*

$$x_i = c_i + \sum_{j=1}^{N} x_{ij} \quad \forall i \in \{1,\ldots,N\} \quad and \quad \sum_{i=1}^{N} \ell_i = 1.$$

We then have the following result:

Proposition 2.26 *The economy admits a competitive equilibrium with quantities* $\mathbf{c}^*$, $\mathbf{x}^*$, $\mathbf{X}^*$, *and* $\boldsymbol{\ell}^*$, *and prices* $\mathbf{p}^*$ *and* w^*.

We define $y_i^* = p_i^* x_i^*$ as the nominal value of output in sector i and $\mathbf{y}^*$ as the vector of nominal values of output across all N sectors. We are interested in the nominal value of aggregate output, y_{agg}:

$$y_{agg} = \mathbf{1}^T \mathbf{y}^* = \sum_{i=1}^{N} p_i^* x_i^*.$$

In the next theoretical result, we compute $\mathbf{y}^*$ in closed form, from which we can then compute in closed form the nominal value of aggregate output, otherwise known as GDP: $y_{agg} = \mathbf{1}^T \mathbf{y}^*$.

Proposition 2.27 *The vector of equilibrium levels of nominal output for all N sectors is:* $\mathbf{y}^* = (\mathbf{I} - \boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1} \boldsymbol{\eta} w$. *The nominal value of aggregate output is* $y_{agg} = \mathbf{1}^T \mathbf{y}^*$.

With $\boldsymbol{\Lambda}$ column-stochastic and $\beta_i \in (0, 1)$ for every sector i, $\mathbf{I} - \boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta})$ is invertible[6] and GDP is readily computable. We can re-write $\mathbf{y}^*$ as follows:

$$\mathbf{y}^* = \left(\sum_{j=0}^{\infty} (\boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta}))^j \right) \boldsymbol{\eta} w = \left(\mathbf{I} + \boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta}) + (\boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta}))^2 + \cdots \right) \boldsymbol{\eta} w.^{7}$$

From this expansion, we see that aggregate expenditure in a particular sector depends on both demand from the representative consumer, the first term in the expansion, and industry demand from other sectors, the remaining terms in the expansion. Industry demand for a particular sector's good comes from both first-order and higher-order connections in the production network. Industry demand from first-order connections arises when there are sectors directly requiring that particular good as an intermediate input. Meanwhile, industry demand arises from higher-order connections when our good of interest indirectly appears in the output of another sector via a supply chain that is greater than length 1.

We would like to study what happens to GDP when a group of sectors receives a positive demand shock. To determine this effect, we first compute the baseline level of GDP in the absence of any transfers or stimulus, y_{agg}^{no}. We then compare it to the level of GDP following the implementation of a particular policy that adjusts the final demand of various sectors. The baseline level of GDP is as

[6] The invertibility of $\mathbf{I} - \boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta})$ is proven in Lemma B.2, which can be found in Appendix B.1.

[7] This expansion holds provided that 1 is smaller than the norm of the inverse of the largest eigenvalue of $\boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta})$. As demonstrated by Lemma B.2 in Appendix B.1, $r(\boldsymbol{\Lambda} \,\mathrm{diag}\,(\boldsymbol{\beta})) < 1$, so this condition holds.

follows:

$$y_{agg}^{no} = \mathbf{1}^T \mathbf{y}^* = N \left[\mathbf{d}_w^- \left((\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right) \right]^T \boldsymbol{\eta} w,$$

where $\mathbf{d}_w^- \left((\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right) = \frac{1}{N} \left[(\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right]^T \mathbf{1}$ is the vector of average weighted in-degrees for the network $\mathcal{G}\left((\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right)$. Agents' weights sum to k: $\mathbf{1}^T \mathbf{d}_w^- \left((\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right) = k$. In our environment, we have two relevant networks: (1) the production network, $\mathcal{G}(\mathbf{Z}') = \mathcal{G}(\boldsymbol{\Lambda})$, that captures the flow of intermediate goods, and (2) the network, $\mathcal{G}(\mathbf{Z}) = \mathcal{G}(\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1}$, that determines each sector's effective weight in the production ecosystem. Sectors' weights, $\mathbf{d}_w^- \left((\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right)$, determine how much of an effect a demand shock has on overall GDP. The more influential the sector, the greater the effect on GDP.

We study two different settings in which a group of sectors receives a positive demand shock: (1) a setting with transfers and (2) a setting with stimulus. For the setting with transfers, a group of sectors receives a positive shock to final demand, while all other sectors receive a negative demand shock. For the setting with stimulus, a group of sectors receives a positive shock to final demand, while all other sectors receive no shock. We are interested in the level of GDP and the corresponding economic multiplier in these two settings. The economic multiplier captures the change in GDP when a group of sectors receives a positive shock to final demand that is of some particular magnitude.

To show how the nominal value of aggregate output changes in settings with transfers and stimulus, we first rewrite y_{agg}^{no}:

$$y_{agg}^{no} = N \left[\mathbf{d}_w^- \left((\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}))^{-1} \right) \right]^T \boldsymbol{\omega},$$

where $\boldsymbol{\omega} = \boldsymbol{\eta} w$. The vector $\boldsymbol{\omega}$ is a vector of sector-specific expenditures on final goods. These expenditures are made by the representative consumer: $p_i^* c_i^* = [\boldsymbol{\omega}]_i$.

In a setting with transfers or externally funded stimulus, the vector of expenditures on final goods across the N sectors changes from $\boldsymbol{\omega}$ to $\boldsymbol{\omega} + \boldsymbol{\rho}$.

Specifically, in a setting with transfers, $[\boldsymbol{\rho}]_i = \epsilon$ if sector i is receiving a positive shock to final demand and $[\boldsymbol{\rho}]_i = -\frac{n\epsilon}{N-n}$ if sector i is not receiving a positive shock to final demand. In this setting with transfers, the total expenditure on final goods remains unchanged from its no-intervention level: $\mathbf{1}^T(\boldsymbol{\omega} + \boldsymbol{\rho}) = \mathbf{1}^T\boldsymbol{\omega}$. We now describe how the transfer of funds across sectors gets implemented. We have a government that is interested in making purchases. Let $p_i^* g_i^*$ be the amount of government expenditure in sector i, where g_i^* is the total number of units of good i purchased by the government. To finance its purchases, the government levies a lump-sum tax τ on the representative consumer. The amount of expenditure by the representative consumer in every sector i is then $p_i^* c_i^* = \eta_i (w - \tau)$. The amount of expenditure by the government is $p_i^* g_i^* = \eta_i \tau + \epsilon$ if sector i is receiving a positive transfer and $p_i^* g_i^* = \eta_i \tau - \frac{n\epsilon}{N-n}$ if sector i is not receiving a positive transfer. n sectors, in total, receive a positive shock to final demand. Note that the total amount of expenditure by the government is equal to its tax revenue: $\sum_{i=1}^N p_i^* g_i^* = \tau$. Also note that $p_i^* c_i^* + p_i^* g_i^* = [\boldsymbol{\omega}]_i + [\boldsymbol{\rho}]_i$ for every sector $i \in \{1, \ldots, N\}$.

In a setting with stimulus, $[\boldsymbol{\rho}]_i = \epsilon$ if sector i is receiving a positive shock to final demand and otherwise $[\boldsymbol{\rho}]_i = 0$. Now, the total expenditure on final goods increases by $n\epsilon$ units relative to the no-intervention level: $\mathbf{1}^T(\boldsymbol{\omega} + \boldsymbol{\rho}) = \mathbf{1}^T\boldsymbol{\omega} + n\epsilon$. The government receives these $n\epsilon$ units of wealth from an external source. The government sets $p_i^* g_i^* = \epsilon$ if it wishes to provide positive stimulus to sector i, and otherwise $p_i^* g_i^* = 0$. n sectors receive positive stimulus. Note that $\sum_{i=1}^N p_i^* g_i^* = n\epsilon$ and that $p_i^* c_i^* + p_i^* g_i^* = [\boldsymbol{\omega}]_i + [\boldsymbol{\rho}]_i$ for every sector $i \in \{1, \ldots, N\}$.

Depending on which sectors get a positive shock to final demand, we can have variation in the resulting level of GDP and the corresponding economic multiplier.

Configuration vector $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ identifies which subset of $n \leq N$ sectors is receiving a positive demand shock. Element $b_i = 1$ if sector i receives a positive shock to final demand and otherwise $b_i = 0$.

We begin by characterizing GDP and the economic multiplier on GDP in a setting with transfers. Sectors $1,\ldots,n$ receive a positive shock to final demand and sectors $n+1,\ldots,N$ receive a negative shock to final demand so that there is zero net transfer across sectors. In this setting, $b_i = 1$ for $i \in \{1,\ldots,n\}$:

$$y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1},\mathbf{b},N,n,0\right) = y_{agg}^{no} + N\epsilon\left[\left(\left[\mathbf{d}_w^{-}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1}\right)\right]_1 + \cdots + \left[\mathbf{d}_w^{-}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1}\right)\right]_n\right) - \frac{n}{N-n}\left(\left[\mathbf{d}_w^{-}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1}\right)\right]_{n+1} + \cdots + \left[\mathbf{d}_w^{-}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1}\right)\right]_N\right)\right]$$

with $m\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1},\mathbf{b},N,n,0\right) = \frac{dy_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1},\mathbf{b},N,n,0\right)}{d\epsilon}$. We continue by characterizing GDP and the economic multiplier on GDP in a setting with stimulus. Now, sectors $1,\ldots,n$ receive a positive shock to final demand while sectors $n+1,\ldots,N$ receive zero shock to final demand. Setting $b_i = 1$ for $i \in \{1,\ldots,n\}$, we have

$$y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1},\mathbf{b},N,n,1\right) = y_{agg}^{no} + N\epsilon\left(\left[\mathbf{d}_w^{-}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1}\right)\right]_1 + \cdots + \left[\mathbf{d}_w^{-}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1}\right)\right]_n\right)$$

with $m\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1},\mathbf{b},N,n,1\right) = \frac{dy_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}(\boldsymbol{\beta}))^{-1},\mathbf{b},N,n,1\right)}{d\epsilon}$.

We are interested in all possible levels of GDP and all possible multipliers given that n sectors receive a positive shock to final demand. We therefore introduce random variables that allow us to characterize the distribution of possible levels of GDP as well as the distribution of possible economic multipliers given n:

Proposition 2.28 *In a setting with transfers, the GDP and the corresponding economic* *multiplier are:*

$$Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,0\right) = y_{agg}^{no} + \frac{N^2\epsilon}{N-n}\left[\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n\right)-\frac{kn}{N}\right] \text{ and}$$

$$M\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,0\right) = \frac{N^2}{N-n}\left[\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n\right)-\frac{kn}{N}\right].$$

In a setting with stimulus, the GDP and the corresponding economic multiplier are:

$$Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right) = y_{agg}^{no} + N\epsilon\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n\right) \text{ and}$$
$$M\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right) = N\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1}\right).$$

The expressions for $Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,0\right)$, $M\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,0\right)$, $Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right)$, and $M\left((\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right)$ in Proposition 2.28 directly map to Equations 2.2-2.5 in Section 2.2, where we set $\mathbf{Z}=(\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\,(\boldsymbol{\beta}))^{-1}$ and $\gamma_1=1$.

All of the theoretical results from Section 2.2 therefore map to the present environment. Given that we are in a setting with transfers or a setting with stimulus and n sectors have received a positive shock to final demand, we can directly compute in closed-form the resulting mean level of GDP and the mean economic multiplier. We can compute the variance of these distributions as well. Moreover, for a particular production network $\mathcal{G}(\mathbf{\Lambda})$, we can determine the lowest possible level of GDP, the highest possible level of GDP, the lowest possible economic multiplier, and the highest possible economic multiplier given that n sectors are receiving a positive shock to final demand. For all feasible production networks $\mathcal{G}(\mathbf{\Lambda})$ with N sectors, for any given N, and n sectors receiving a positive shock to final demand, we can approximate the cumulative distribution functions for GDP and the corresponding economic multiplier. Quite importantly, we can analytically determine the probability that a particular policy measure, whether it be a transfer of funds across sectors or externally funded stimulus, leads to a reduction in GDP below its no-intervention level and a negative economic multiplier. GDP dips below its no-intervention level when the economic multiplier is negative.

In settings with stimulus, it turns out that the probability of a negative multiplier is always zero. GDP following an intervention that targets n sectors for stimulus is at least as large as GDP in the absence of any intervention. This property holds for all feasible levels of stimulus, that is, for all $n \in \{1, \ldots, N-1\}$:

Proposition 2.29 *For every $n \in \{1, \ldots, N-1\}$,*

$$\Pr\left[Y_{agg}\left((\mathbf{I} - \mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}, N, n, 1\right) \geq y_{agg}^{no}\right] = 1 \text{ and}$$
$$\Pr\left[M\left((\mathbf{I} - \mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}, N, n, 1\right) \geq 0\right] = 1.$$

Stimulus always weakly increases GDP. Given ϵ units of stimulus for n sectors, the multiplier is always non-negative.

The results that we have provided thus far hold for any feasible network structure $\mathcal{G}(\mathbf{\Lambda})$. We now study what happens to the distribution of GDP and the distribution of corresponding economic multipliers when the network is trivial. Specifically, we consider an environment in which there is no network-based interaction. To be consistent with the other sections, we declare an absence of network-based interaction when $\mathbf{\Lambda} = \mathbf{I}$, the identity matrix:

Proposition 2.30 *In the absence of any network-based interaction,* $Y_{agg}\left((\mathbf{I} - \mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}, N, n, 0\right)$, $M\left((\mathbf{I} - \mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}, N, n, 0\right)$, $Y_{agg}\left((\mathbf{I} - \mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}, N, n, 1\right)$, *and* $M\left((\mathbf{I} - \mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}, N, n, 1\right)$ *have the same* *functional form as in Proposition 2.28, with*

$$\left[\mathbf{d}_w^- \left(\left(\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}) \right)^{-1} \right) \right]^T = \frac{1}{N} \begin{pmatrix} \frac{1}{1-\beta_1} & \cdots & \frac{1}{1-\beta_N} \end{pmatrix}.$$

Provided that we do not have $\beta_1 = \cdots = \beta_N$, the distributions of GDP and economic multipliers are non-degenerate. Both GDP and the corresponding economic multiplier can vary with the particular configuration of transfers or stimulus even when there is no underlying network structure. This feature of non-degeneracy distinguishes the present environment from those of the other two sections. In the other two sections, once we removed the network structure, the distribution of the aggregate action and the distribution of the corresponding economic multiplier both became degenerate. In those two environments, by removing the agent interaction structure, all agents became identical; the aggregate action and the economic multiplier were the same regardless of which subset of agents received a positive transfer or positive stimulus. In the present environment with production, even though we are eliminating any heterogeneity that arises from the topology of the production network, there is still heterogeneity across sectors; each sector differs in the intensity with which it uses intermediate inputs, which makes the level of GDP vary and the value of the economic multiplier vary with the particular configuration of transfers or stimulus.

Once $\beta_1 = \cdots = \beta_N \equiv \beta$, both GDP and the corresponding economic multiplier are invariant to configuration for all levels, n, of transfers or stimulus:

Proposition 2.31 *Both GDP and the corresponding economic multiplier are invariant to configuration if and only if $\beta_1 = \cdots = \beta_N \equiv \beta$:*

$$Y_{agg}\left(\left(\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}) \right)^{-1}, N, n, 0 \right) = y_{agg}^{no} \quad \textit{and} \quad M\left(\left(\mathbf{I} - \boldsymbol{\Lambda} \operatorname{diag}(\boldsymbol{\beta}) \right)^{-1}, N, n, 0 \right) = 0$$

with probability 1, and

$$Y_{agg}\left(\left(\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\left(\boldsymbol{\beta}\right)\right)^{-1},N,n,1\right)=y^{no}_{agg}+\frac{n\epsilon}{1-\beta}\ \ \textit{and}$$

$$M\left(\left(\mathbf{I}-\mathbf{\Lambda}\,\text{diag}\left(\boldsymbol{\beta}\right)\right)^{-1},N,n,1\right)=\frac{n}{1-\beta}$$

with probability 1.

When every sector has the same labor intensity, that is, $\beta_1 = \cdots = \beta_N$, both GDP and the corresponding multiplier are invariant to configuration for any production network $\mathcal{G}(\mathbf{\Lambda})$, provided that $\mathbf{\Lambda}$ is column-stochastic. In such an environment, the distribution of GDP and the distribution of possible economic multipliers are both degenerate. In settings with transfers, GDP equals its no-intervention level regardless of which group of sectors receives a positive shock and regardless of how many sectors receive this positive shock. This result holds for all feasible networks $\mathcal{G}(\mathbf{\Lambda})$. In settings with stimulus, there is an increase in GDP relative to its no-intervention level, but this level of GDP remains the same regardless of which group of sectors receives positive stimulus.

2.6 Conclusion

This paper studies economic systems with N networked agents, $n \leq N$ of whom each initially receive a positive shock that is either financed by internal transfers or external stimulus. This work examines the resulting probability distributions of possible aggregate actions and economic multipliers given n. Agents' actions are interconnected, with such interdependency captured by the topology of an interaction network. As the particular configuration of positive shocks to agents varies, holding n fixed, agents' actions adjust so that the aggregate action varies as well. In general, the distribution of possible aggregate actions and the distribution of possible economic multipliers are non-degenerate. We explore these distributions of aggregate actions and corresponding economic multipliers in three different networked environments: (1) those featuring strategic complements and substitutes, (2) those featuring coordination and anti-coordination, and (3) those featuring production.

Despite strong differences across these three environments, the core mathematics is the same. For each environment, given agents' decision-making behavior and the underlying network structure, we construct a network-derived vector of agent weights, the vector of average weighted in-degrees for graph $\mathcal{G}(\mathbf{Z})$. It is from this vector of agent weights that we can characterize in closed form all essential features of the distribution of aggregate actions and the corresponding distribution of economic multipliers. For all feasible population sizes N, number of agents receiving a positive shock, n, and network topologies, we can compute the mean aggregate action and the mean economic multiplier. We can also compute in closed form the variance for these two distributions, the bounds on the support of these distributions, the corresponding limiting distributions as $N \to \infty$, and approximations to the CDFs of these two distributions for finite N. We study the aggregate action and the economic multiplier when there is no underlying network structure and when the network topology is such that these two quantities are invariant to configuration. In addition, for certain environments, we can rank networks so that the distributions of possible aggregate actions and economic multipliers corresponding to a higher-ranked network first-order stochastically dominate the distributions of possible aggregate actions and economic multipliers for a lower-ranked network. As a result, the higher-ranked the network, the more effective the policy.

Quite importantly, we develop and use a set of tools that allows us to analytically compute the probability of a negative economic multiplier given agents' decision-making behavior and a particular aggregate level, $n\epsilon$, of transfers or stimulus. The set of network-derived agent weights ultimately shapes the probability that a particular environment generates negative economic multipliers. In settings with transfers, practically every network structure generates a negative multiplier with a positive probability. Quantifying the probability of a negative economic multiplier or an aggregate action below its no-intervention level is important because it captures the extent to which a policy is ineffective. Policy-making entities craft policy measures to achieve particular objectives, such as jump-starting an economy during recession. If the naturally occurring agent interaction structure is such that there is a non-negligible positive probability of a negative economic multiplier, then the policy-making entity may rethink its policy prescription. When the policy is financed by internal transfers, there is no outright cost to the policy-making entity, except that it is transferring funds away from individuals. However, when the policy is financed by external stimulus, perhaps by issuance of debt, the policy-making entity ultimately must provide repayment, and if the economic multiplier following stimulus is negative, it becomes all the more difficult to repay borrowed funds. If the policy-making entity does not have a good grasp of the topology of agents' interaction structure, then enacting a policy can be quite risky, as there are entire non-degenerate distributions of possible resulting aggregate actions and economic multipliers.

In essence, this work extends and applies a set of tools that allows us to construct in closed form policy-induced distributions of possible aggregate actions and economic multipliers in environments with complex, network-based agent interactions. The tools and the methodology implemented in this work are general. Hopefully they can be applied to a broad range of diverse settings and provide substantive insights.

Chapter 3

Comprehensively Stress Testing the Economy

3.1 Introduction

The 2008 global financial crisis and the concomitant Great Recession consisted of an unprecedented series of events and adjustments to the macroeconomic and financial landscape. Within the domestic housing market, home prices fell approximately 30 percent from mid-2006 to mid-2009.[1] In the United States, both investment bank Lehman Brothers and savings and loan institution Washington Mutual failed. Financial institutions Bear Stearns, Merrill Lynch, AIG, Freddie Mac, Fannie Mae, and Wachovia experienced some form of rescue or bailout. Globally, Northern Rock, HBOS, Royal Bank of Scotland, Bradford & Bingley, Fortis, Hypo Real Estate, and Alliance & Leicester likewise experienced some form of rescue, bailout, and/or

[1] https://www.federalreservehistory.org/essays/great_recession_of_200709. Accessed March 18, 2019.

nationalization.[2] The unemployment rate in the United States increased from 5 percent in December 2007 to 10 percent in October 2009, real United States GDP contracted by 4.3 percent between the fourth quarter of 2007 and the second quarter of 2009, and the S&P 500 index fell 57 percent from its peak in October 2007 to its trough in March 2009. Net worth across United States households and nonprofit organizations fell from a peak of approximately 69 trillion dollars in 2007 to a trough of approximately 55 trillion dollars in 2009.[3] To bring greater stability to the financial system, the United States federal government implemented the Troubled Asset Relief Program (TARP), in which it purchased troubled companies' assets and equity. For TARP, Congress authorized the United States Treasury 475 billion dollars to make purchases; the Treasury principally used this money to stabilize banks, develop programs to increase credit availability, rescue the United States automobile industry, stabilize AIG, and buttress programs that prevent foreclosure.[4] Globally, real GDP growth decreased from 5.6 percent in 2007 to 0.1 percent in 2009[5], and the financial crisis spurred a European sovereign debt crisis for the countries of Iceland, Portugal, Italy, Ireland, Greece, Spain, and Cyprus.

The global financial crisis led different governmental and supervisory authorities to massively reassess both their regulatory roles and existing financial regulation. The Basel Committee on Banking Supervision developed a set of recommendations for regulation, known as Basel III, in response to the global financial crisis. Basel III

[2] https://www.theguardian.com/business/2008/dec/28/markets-credit-crunch-banking-2008. Accessed March 18, 2019.

[3] https://www.federalreservehistory.org/essays/great_recession_of_200709. Accessed March 18, 2019.

[4] https://www.investopedia.com/terms/t/troubled-asset-relief-program-tarp.asp. Accessed March 18, 2019.

[5] https://blog.euromonitor.com/global-economy-10-years-after-financial-crisis. Accessed March 18, 2019.

focused on strengthening regulation, supervision, and risk management of financial institutions to ensure financial stability. For example, Basel III sought to improve the quality of bank regulatory capital, increase the level of capital requirements, constrain excess leverage, and mitigate excess liquidity risk. The Basel Committee included representatives from central banks and regulatory authorities around the world, and in general, its members applied these standards in their own jurisdictions; indeed, the Federal Reserve announced that in December 2011 it would implement essentially all of the Basel III regulations. Within the United States, the Dodd-Frank Wall Street Reform and Consumer Protection Act, signed into federal law in July 2010, represented the federal government's response to the global financial crisis. The Dodd-Frank Act made significant changes to financial regulation to improve financial stability and consumer protection. It promoted financial stability through creation of the Financial Stability Oversight Council and the Office of Financial Research, and it advocated for consumer protection in the financial industry through creation of the Consumer Financial Protection Bureau. The Dodd-Frank Act additionally introduced corporate governance reforms, executive compensation reforms, credit rating agency regulation, securitization retention requirements, procedures for regulatory enforcement, and regulation of over-the-counter derivatives, among other types of regulation.[6]

Both Basel III and the Dodd-Frank Act led to major adjustments in financial regulation. They equipped regulators and supervisory institutions with new tools to assess the stability of individual financial institutions and the financial system as a whole, and they broadened these supervisory institutions' mandates. For example, Basel III introduced a set of statistics for central banks to collect from

[6] https://corpgov.law.harvard.edu/2010/07/07/summary-of-dodd-frank-financial-regulation-legislation/. Accessed March 18, 2019.

the balance sheets of financial institutions. Basel III provided minimum and/or maximum allowable values for these statistics, which included capital ratios, and failure to meet these guidelines forced financial institutions to adjust their balance sheets. Given all of these new regulations and guidelines, we might wonder whether the resulting constrained financial system would now be able to survive and maintain normal operations after encountering the same economic conditions that had precipitated the 2008 global financial crisis. To answer this question, we would need to construct a scenario mimicking the start of the financial crisis and the Great Recession, and we would need to examine whether the financial system could withstand such stresses. This process of scenario design and stress testing is exactly what the Federal Reserve decided to do in response to the global financial crisis. A massive part of the Federal Reserve's post-crisis regulatory toolkit involved designing stressful economic and financial scenarios and quantifying these scenarios' effects on financial institutions' capital holdings and balance sheets.

Following the financial crisis, in May 2009, the committee behind Basel III published guidelines for stress testing. Within the United States, the Dodd-Frank Act mandated that the Federal Reserve conduct annual supervisory stress tests for sufficiently large bank holding companies, which essentially includes large banks and other large financial institutions. The Federal Reserve decided to jointly implement these Dodd-Frank Act stress tests and a process of comprehensive capital analysis and review (CCAR); the former focused on studying financial institutions' balance sheet items under stressed scenarios, while the latter focused on studying financial institutions' capital adequacy under stressed scenarios, thus making the Federal Reserve's annual supervisory stress tests and CCAR complementary. Not only is stress testing a core part of the supervisory toolkit, it is also a key tool for internal risk management by financial institutions. The Federal Reserve's stress tests, in principal, can potentially capture broad risks, while individual financial institutions' stress tests can capture the idiosyncratic risks that are germane for each institution.

Adhering to the requirements of the Dodd-Frank Act, the Federal Reserve undertakes the following stress testing process. It conducts annual supervisory stress tests for sufficiently large financial institutions. Each stress test involves crafting a different real-world scenario that perturbs certain economic quantities of interest. With the Federal Reserve required to carry out three stress tests, it therefore crafts three different scenarios: (1) a baseline scenario, (2) an adverse scenario, and (3) a severely adverse scenario. The Federal Reserve uses data provided by the financial institutions, and it examines how the three different stress scenarios impact each financial institutions' balance sheets and capital holdings. There are many possible economic variables to perturb in constructing stress tests, so the Federal Reserve takes the following approach. It requires every sufficiently large financial institution to undertake a macroeconomic stress test. The baseline scenario, the adverse scenario, and the severely adverse scenario are all constructed from the perturbations of mainly macroeconomic variables, such as the unemployment rate or the growth rate of GDP. Financial institutions that have significant trading activity must also add a global market shock to their macroeconomic stress tests. The global market shock is an add-on component to the macroeconomic adverse scenario and the severely adverse macroeconomic scenario. For the global market shock, various market factors can adjust, such as equity prices, private equity values, and foreign exchange rates. Financial institutions with substantial trading or processing and custodian operations must additionally incorporate counterparty default into their stress tests. Counterparty default is an add-on component to each stress test. Each financial institution must estimate and report potential losses and effects on capital that would result if the institution's largest counterparty unexpectedly defaulted. The Federal Reserve applies the same set of supervisory stress tests to each financial institution. In addition to these supervisory scenarios, every financial institution must carry out its own internal stress test. The results of the Federal Reserve's supervisory stress tests and each financial institution's stress test are publicly disclosed.

The Federal Reserve's stress testing approach has several limitations. The number of stress tests that get conducted annually is very small. Each year, the Federal Reserve only carries out three supervisory stress tests for every sufficiently large financial institution: a baseline scenario, an adverse scenario, and a severely adverse scenario. Each supervised financial institution must also execute one additional stress test. These four scenarios are not enough to ensure that the financial system has been comprehensively stress tested. Indeed, the Federal Reserve's baseline scenario, adverse scenario, and severely adverse scenario sometimes involve perturbations of the same underlying economic variables. What then distinguishes these scenarios are the magnitudes by which these perturbed economic variables adjust; the severely adverse scenario, for example, might simply feature greater magnitudes of adjustment for the underlying economic variables than the adverse scenario. As a result, the supervisory scenarios do not necessarily capture different ways that a financial system can become stressed. Moreover, the stress test scenarios crafted by individual financial institutions are sometimes inspired by the Federal Reserve's supervisory stress tests; consequently, these scenarios do not identify the idiosyncratic risks that can potentially destabilize individual financial institutions.

There exist additional weaknesses beyond the number of annual stress tests being very small. The supervisory stress test scenarios are often calibrated to past historical events. While that does serve as a reasonable benchmark, it is unlikely that history will exactly repeat itself. The economy may inevitably evolve along certain pathways that precipitate recessions and/or financial crises, but it is unlikely that these pathways will always be identical; the circumstances generating recessions and/or financial crises are not always the same. Financial institutions might be able to withstand stress test scenarios that mimic the onset of the 2008 global financial crisis, but that does not mean that the financial system is stable. The Federal Reserve may as a result be lulled into a false sense of financial system stability. Moreover, the Federal Reserve's stress test scenarios have not substantially changed over the years. With financial institutions adjusting their operations and portfolios to satisfy the Federal Reserve's stress tests, they may all become vulnerable to other realistic forms of risk. It is difficult to know what exactly should be the ideal set of stress test scenarios to ensure financial system stability. For instance, the Federal Reserve's current approach to stress testing generally assumes that stress originates within the macroeconomy, hence its stress testing scenarios primarily being macroeconomic scenarios. It is, however, plausible that stresses within the financial system initiate financial crises, which would make market risk scenarios more relevant.

The current stress testing approach within the United States presents several legitimate concerns. This chapter devises a solution to address all of these concerns. We can tackle all of the stated weaknesses in the Federal Reserve's current stress testing approach by massively increasing the number of distinct stress tests conducted annually. One might argue that drastically increasing the number of distinct stress test scenarios generates its own set of problems: for example, it can enormously increase the computational burden, and the process of generating additional scenarios can potentially be very haphazard. The present work addresses these critiques. It develops a systematic approach for scaling up the number of stress test scenarios, and it shows how to substantially increase the number of stress tests without increasing the computational burden. The procedure presented in this work for comprehensively stress testing the economy and the financial system can benefit the Federal Reserve, it can benefit individual financial institutions that carry out internal assessments of risk, and it can benefit central banks and supervisory institutions globally.

To illustrate the approach that the present work takes, let's start off with an example. Imagine that there exists a stress test scenario in which the Indian rupee depreciates by 10 percent. We are interested in the effects of this depreciation on the balance sheets of individual financial institutions and the financial system as a whole. Given this particular stress test scenario, we can generate an entire class of stress tests. At the highest level, what we are interested in here is the effect of exchange rate risk on the financial system. We therefore develop a class of stress tests, and each stress test within this class is distinguished by the type of foreign currency facing a 10-percent depreciation; the first stress tests features a 10-percent depreciation of the Indian rupee, the second stress test features a 10-percent depreciation of the euro, the third stress test features a 10-percent depreciation of the Mexican peso, and so on. The number of stress tests within this class is then equal to the number of foreign currencies. Therefore, given our one initial stress test, we have generated an entire class of possible stress tests. For this class, we can construct in closed form a probability distribution that summarizes balance sheet effects for each individual financial institution, and we can construct a probability distribution that summarizes balance sheet effects for the entire financial system. Rather than having individual data points, we have entire probability distributions capturing exchange rate risk for the financial system.

Now, let's instead imagine that there exists a stress test scenario in which the Indian rupee depreciates by 10 percent and the euro appreciates by 15 percent.

Here, we have a different form of exchange rate risk. As in the previous example, we proceed to generate an entire class of stress test scenarios capturing this form of exchange rate risk. For every stress test within the class, we pick one currency to depreciate by 10 percent, and we pick a separate currency to appreciate by 15 percent. The class exhausts all possible combinations of appreciating and depreciating currencies, so that the total number of stress tests within this class is combinatorial. Given this class of stress tests, there is a corresponding probability distribution summarizing balance sheet effects for each financial institution, and there is a corresponding probability distribution summarizing balance sheet effects for the overall financial system; we can solve for the major statistical features of these probability distributions in closed form.

The present work thus takes the following approach. It constructs different classes of stress tests. Each of these classes of stress tests is distinguished by its categories of risk. For example, one class of stress tests might feature certain levels of exchange rate risk and sectoral risk, while another class of stress tests might feature a particular magnitude increase in the probability of default for a type of debt. There are different ways that these types of risk can enter into the financial system; for example, exchange rate risk can separately manifest itself in assets denominated in different currencies. Each stress test scenario within a particular class therefore represents a different way that these types of risk manifest themselves. The set of stress test scenarios within a particular class is exhaustive; there are no additional ways that these types of stresses can be distributed within the financial system. For each class of stress tests, we construct a probability distribution that captures balance sheet effects for each individual financial institution, and we construct a probability distribution that captures balance sheet effects for the overall financial system. Through this approach, we massively increase the number of stress test scenarios in a systematic fashion without increasing the computational burden.

One last criticism of the Federal Reserve's current stress testing approach is that it is not sufficiently macroprudential. Even though the global financial crisis exposed the weaknesses of regulatory approaches that are too microprudential, the Federal Reserve's shift from a microprudential regulatory approach to one that is relatively more macroprudential has mostly been nominal. The present work takes steps towards making the Federal Reserve's regulatory approach relatively more macroprudential. Macroprudential regulation is concerned with risks at the level of the financial system. The present work approaches the stress testing process by identifying categories of risk and then specifying the different ways that such risk can manifest itself within the financial system. This top-down perspective is fundamentally macroprudential. The present work moreover discerns how networks, a fundamentally macroprudential object, shape financial stability. One of the main networks studied in this work is a bipartite network that links financial institutions to assets. The present work shows, given a particular class of stress tests, how the topology of the bipartite network shapes stress tests' effects on the financial system. More precisely, the topology of the bipartite network determines the shape of this corresponding probability distribution.

3.1.1 Relation to the Literature

The present chapter focuses on stress tests and how to massively improve this major part of the Federal Reserve's supervisory toolkit. Hirtle and Lehnert (2015) provides background on stress testing in the United States and discusses the objectives of stress testing. Glasserman and Tangirala (2016), Demekas (2015), and Anderson (2016) also provide a history and overview of stress tests. Acharya et al. (2014) and Borio et al. (2014) discuss macroeconomic stress tests, in which macroeconomic factors adjust and cause the transmission of shocks to the financial system. A weakness of macroeconomic stress tests, as Bookstaber et al. (2014) discuss, is that they neglect scenarios in which shocks to the financial system themselves cause economic downturns. Consistent with this critique, the present work considers finance-specific stress tests scenarios that directly affect financial institutions' balance sheets rather than scenarios primarily motivated by macroeconomic adjustments. Petrella and Resti (2013) and Scheurmann (2014) both weigh the costs and benefits of publicly disclosing the results of stress tests.

The literature has highlighted a couple of weaknesses regarding the Federal Reserve's stress testing approach. First, as Glasserman and Tangirala (2016) mention, the Federal Reserve's stress tests have not drastically changed over the years. Financial institutions have adjusted their balance sheets so that they can withstand the Federal Reserve's stress tests, but in doing so, they may not be able to survive an actual stress scenario that differs from the ones implemented by the Federal Reserve. Second, the Federal Reserve carries out too few stress test scenarios; Bookstaber et al. (2014), Demekas (2015), and Glasserman and Tangirala (2016) all discuss the need to increase the number of stress test scenarios. Grundke (2011) addresses the issue of scenario selection by instead carrying out reverse stress tests; instead of deciding which scenarios are appropriate, the supervisory institution identifies a certain outcome or threshold of interest and then generates stress scenarios that would yield that particular outcome. Reverse stress tests have their own weaknesses. For reverse stress tests, the outcome or threshold must be very specific, and the number of stress test scenarios that can yield that particular outcome is potentially extremely large. It is often not feasible for the regulatory institution to entertain all of these possible stress scenarios. The present work instead focuses on massively increasing the number of stress test scenarios without increasing the computational burden. It generates classes of stress tests that each contain a very large number of individual stress tests. Associated with each class of stress tests is a corresponding probability distribution that summarizes the effects of those stress tests.

The present work interfaces with the literature on macroprudential regulation; it shows how to design stress tests so that they are more macroprudential in nature, rather than being strictly microprudential. Clement (2010) provides a history of the term "macroprudential," and how it has been used over time. "Macroprudential" is traced back to 1979, in which it was mentioned at a meeting of the Cooke Committee, the predecessor of the Basel Committee on Banking Supervision. At the time, "macroprudential" meant "an enhanced focus on the financial system as a whole and its link to the macroeconomy." Clark and Large (2011), Liebeg and Posch (2011), and Claessens (2015) all provide a modern overview of macroprudential regulation and its objectives. Macroprudential regulation is often concerned with risks at the level of the financial system; it is distinguished from microprudential regulation, which instead seeks to ensure the soundness of individual financial institutions one at a time. Hanson et al. (2011), Kashyap et al. (2011), and Borchgrevink et al. (2014) argue that macroprudential regulation arises out of a need to address market failures. Pecuniary externalities, such as fire sales of assets, interconnectedness externalities, and strategic complementarities all motivate macroprudential regulation because the standard microprudential toolkit does not address these market failures. Acharya (2009) discusses the role of capital requirements in a macroprudential framework. Borio (2003), Greenlaw et al. (2012), and Williams (2015) acknowledge that stress tests and current regulatory frameworks are still fairly microprudential in nature. Williams (2015) argues that microprudential regulations and supervision are unfortunately being used to attain macroprudential objectives due to the scarcity of explicitly macroprudential tools. The present work shows how to substantially enhance the Federal Reserve's existing stress testing approach to make it much more macroprudential. The present work introduces classes of stress tests, which are distinguished by their types of risks. Each individual stress test within a particular class is then distinguished by the specific ways that these risks manifest themselves within the financial system, whether these risks appear in certain assets or financial institutions. Studying the different ways that risk can be distributed within the financial system, and quantifying the corresponding effects on the financial system is a fundamentally macroprudential perspective. Clement (2010) offers this perspective when discussing the history of the term "macroprudential."

Financial networks and their topologies form an important part of the present work. The present work studies bipartite networks that link individual financial institutions to individual assets; edges are directed from the financial institutions to the assets in their portfolios, with the weight of each edge equal to the number of units of the asset held by the financial institution. Institutions' overlapping portfolios here generate risk. As a result of the 2008 global financial crisis, there is quite a large literature on financial networks. Caccioli et al. (2014) and Levy-Carciente et al. (2015) study bipartite networks linking financial institutions to assets, and Marotta et al. (2015) studies bipartite credit networks linking banks to firms. Gualdi et al. (2016) examines portfolio overlap among financial institutions as a channel for financial contagion. Now, a large part of the financial networks literature is focused on counterparty networks and financial contagion. Allen and Babus (2009), Gai and Kapadia (2010), Battiston et al. (2012), Elliott et al. (2014), and Acemoglu et al. (2015) all study how the actual structure of the counterparty network shapes systemic risk. Afonso et al. (2011) studies the impact of the global financial crisis on edge weights in a counterparty network of financial institutions. Cont et al. (2013) tries to determine the systemic importance of each financial institution in an interbank network. Zawadowski (2013) focuses on a counterparty network with bilateral over-the-counter contracts, and Markose et al. (2012) focuses on the network of credit default swaps within the United States at the time of the financial crisis. For Farboodi (2014), the structure of the financial network is determined endogenously as financial institutions make strategic borrowing and lending decisions. The present work studies how the topologies of financial networks shape stress tests' effects on the financial system. The present work carries out this analysis for entire classes of stress tests, not just individual stress tests.

In addition to interfacing with the overlapping literatures on stress testing, macroprudential regulation, and financial networks, the present work contributes to a new literature on networks and probability distributions in the economy. Both Chapter 1 and Chapter 2 of this dissertation provide a foundation for this literature. Chapter 1 develops a set of theoretical tools for mapping the topology of an economic network to a probability distribution of possible outcomes. Chapter 1 adapts these tools to study locally formed macroeconomic sentiment and how agents' interaction structure shapes the capacity for there to exist non-fundamental swings in aggregate macroeconomic sentiment; Chapter 1 thereby enhances our understanding of animal spirits. Chapter 2 extends the set of theoretical tools from Chapter 1 so that they have broader applicability. As in Chapter 1, Chapter 2 also maps naturally occurring networks in the economy to different probability distributions of interest. Chapter 2, in particular, focuses on the effects that a given policy has on a population's aggregate action when agents are networked and the actions that these agents take are interdependent. Chapter 2 therefore explicitly shows, for any given policy targeting a certain number of networked agents, how the topology of agents' interaction network shapes the corresponding distribution of possible aggregate actions and the corresponding distribution of possible economic multipliers. The present work makes additional methodological and technical advances relative to Chapter 1 and Chapter 2; it further expands the set of tools for mapping networks to probability distributions. The present work studies how the topologies of bipartite networks linking financial institutions to assets shape stress tests' effects on the financial system. Given a particular class of stress tests, the present work shows how to map the topology of the bipartite network to a probability distribution capturing balance sheet effects.

3.1.2 Outline of Chapter

We begin Section 3.2 by introducing notation and definitions. We next consider stress tests that directly shock the portfolios of financial institutions. While we are indeed interested in each stress test's effects on the balance sheets of individual financial institutions and the financial system as a whole, we would like to drastically increase the number of such tests to more rigorously stress the financial system. In this section, we therefore show how to drastically increase the total number of stress tests without increasing the computational burden. We generate classes of stress tests, and for each class, we construct a probability distribution capturing possible balance sheet effects for each individual financial institution, and we construct a probability distribution capturing possible aggregate balance sheet effects for the entire financial system. In the United States, these shocks to financial institutions' portfolios generally constitute the *global market shock component* of the Federal Reserve's stress tests, and they can also constitute the macroeconomic component. We conclude in Section 3.3.

3.2 Classes of Stress Tests and Probability Distributions of Balance Sheet Effects

3.2.1 Notation and Definitions

The cardinality of a set $\mathcal{X}$ is $|\mathcal{X}|$. A *multiset* is an object similar to a set, but it allows for multiple instances of each of its elements. Vector $\mathbf{x}$ is a column vector by default. The i^{th} element of vector $\mathbf{x}$ is $[\mathbf{x}]_i$. The ij^{th} element of matrix $\mathbf{X}$ is $[\mathbf{X}]_{ij}$, the i^{th} row of $\mathbf{X}$ is $[\mathbf{X}]_{i*}$, and the j^{th} column of $\mathbf{X}$ is $[\mathbf{X}]_{*j}$. The column vector whose elements all equal 1 is $\mathbf{1}$. The Hadamard product of matrices $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{X} \circ \mathbf{Y}$, is their element-wise multiplication: $[\mathbf{X} \circ \mathbf{Y}]_{ij} = [\mathbf{X}]_{ij} [\mathbf{Y}]_{ij}$. For permutation matrix $\mathbf{P}$, $\mathbf{PX}$ permutes the rows of $\mathbf{X}$ and $\mathbf{XP}$ permutes the columns of $\mathbf{X}$. The $*$ operator denotes the convolution of two probability distributions. $\mathbb{Z}$ is the set of all integers.

3.2.2 Theoretical Framework

The Federal Reserve's stress tests shock the portfolios of financial institutions through multiple conduits. Each stress test includes multiple categories of risk, including sovereign risk, exchange rate risk, and industry risk. We can take these categories of risk as key parts of economic and financial downturns, and we can examine all of the different ways that these categories of risk manifest themselves within the broader economy and financial system.

These different types of risk affect the values of assets; for instance, they change the prices of assets and they alter income streams. We are interested in stress tests' effects on balance sheet items for individual financial institutions and the financial system as a whole, that is, the collection of financial institutions that comprise the financial system. In this section, we focus on stress tests' effects on the market value of institutions' net assets. Specifically, the market value of a financial institution's net assets is equal to the market value of its assets minus the market value of its liabilities. Assets and liabilities for each financial institution are marked to market. A separate balance sheet item that we could instead look at is net income before taxes. Net income takes into account unrealized and realized gains and losses for securities, so net assets is a reasonable balance sheet item to examine.

There are M financial institutions indexed $1, \ldots, M$ and N total securities indexed $1, \ldots, N$. Each security in a financial institution's portfolio is either an asset or a liability. Define the $M \times N$ matrix $\mathbf{A}$. The ij^{th} element of $\mathbf{A}$ is equal to the number of units of security j held in the portfolio of financial institution i. Each row of $\mathbf{A}$ represents a financial institution's portfolio. The elements of $\mathbf{A}$ can be positive, zero, or negative. $[\mathbf{A}]_{ij} = 0$ means that institution i does not have security j in its portfolio, $[\mathbf{A}]_{ij} > 0$ means that security j is one of institution i's assets, and $[\mathbf{A}]_{ij} < 0$ means that security j is one of institution i's liabilities. The extent to which the rows of $\mathbf{A}$ are similar determines the extent to which the portfolios of different financial institutions are overlapping. It determines the extent to which the financial system exhibits systematic risk. Define $\mathbf{p}$ as the $N \times 1$ vector of securities prices. The market value of institution i's net assets is $[\mathbf{A}]_{i*}\,\mathbf{p}$. The market value of net assets for the entire financial system is $\mathbf{1}^T\mathbf{A}\mathbf{p}$.

In the background, we have this bipartite network linking financial institutions to securities. The weight of each directed edge is equal to the number of units of a particular security in the corresponding financial institution's portfolio. We are interested in how the topology of this bipartite network shapes stress tests' effects on individual financial institutions and the overall financial system.

We want to know how a stress test changes each individual institution's net assets and the net assets for the financial system. In the Federal Reserve's current stress testing approach, a stress test features shocks that cause certain securities' prices to change; we then quantify the overall effect. In the present work, we would like to massively increase the number of annual stress tests being conducted without increasing the computational burden. Therefore, rather than identifying individual stress tests, we identify classes of stress tests. Each class of stress tests is distinguished by its categories of risk. Categories of risk, for example, can include a certain form of exchange rate risk, a certain form of sovereign debt risk, and a certain form of industry risk, all of which affect securities prices. Different classes of stress tests differ in at least one category of risk. Since each category of risk can manifest itself in different ways (for instance, exchange rate risk can manifest itself in different types of currencies), a given class of stress tests can contain many different individual stress tests. The individual stress tests within a particular class account for all of the possible ways that the categories of risk can manifest themselves.

When we consider just one stress test, we have one data point that captures the resulting level of net assets for each individual financial institution, and we have one data point that captures the resulting level of net assets for the whole financial system. In the present work, when we instead consider one class of stress tests, we have a probability distribution that captures possible resulting levels of net assets for each individual financial institution, and we have a probability distribution that captures possible resulting levels of net assets for the whole financial system. In this work, we are moving from data points to entire probability distributions.

To compute how a class of stress tests affects individual institutions' net assets and aggregate net assets, we take the following approach. A class of stress tests features one or more categories of risk. We separately consider each category of risk. For each category of risk, we construct a probability distribution of possible changes in net assets for each individual financial institution, and we construct a probability distribution of possible changes in aggregate net assets for the overall financial system. For each individual financial institution, we then convolve the relevant risk category-specific probability distributions, and the resulting probability distribution captures possible changes in net assets for that class of stress tests with its multiple categories of risk. Similarly, for the overall financial system, we convolve the relevant risk category-specific probability distributions, and the resulting probability distribution captures possible changes in net assets for the entire financial system given that particular class of stress tests with its multiple categories of risk.

We start here by focusing on individual categories of risk. For each category of risk, we select the relevant set of affected securities from the full set of securities, and we cluster the affected securities into various groups as needed. From the original matrix $\mathbf{A}$ and the original price vector $\mathbf{p}$, we therefore generate a new matrix $\bar{\mathbf{A}}$ and a new vector $\bar{\mathbf{p}}$. In general, matrix $\bar{\mathbf{A}}$ captures financial institutions' portfolio holdings for the affected securities, with securities clustered as needed, and vector $\bar{\mathbf{p}}$ captures the initial values of these securities clusters. The price vector $\bar{\mathbf{p}}$ can potentially differ for each financial institution and for the overall financial system; when that occurs, we define $\bar{\mathbf{p}}_i$ as the price vector for financial institution i, and we define $\bar{\mathbf{p}}_{agg}$ as the price vector for the overall financial system. We then introduce the vector $\mathbf{w}_i = \left([\bar{\mathbf{A}}]_{i*}\right)^T$ of risk category-relevant portfolio holdings for each individual financial institution $i \in \{1, \ldots, M\}$, and we similarly construct a vector $\mathbf{w}_{agg}$ of risk category-relevant portfolio holdings for the entire financial system. We will use $\mathbf{w}_i$ and $\mathbf{w}_{agg}$ to respectively compute the effects of a category of risk on net assets for financial institution i and for the overall financial system.

The next three examples illustrate how to construct $\bar{\mathbf{A}}$, $\bar{\mathbf{p}}$ (or $\bar{\mathbf{p}}_i$ and $\bar{\mathbf{p}}_{agg}$ as needed), $\mathbf{w}_i$ for all $i \in \{1, \ldots, M\}$, and $\mathbf{w}_{agg}$ from $\mathbf{A}$ and $\mathbf{p}$ given a particular category of risk:

Example 3.1 *Suppose that the category of risk is credit risk. Specifically, forty percent of all unique AAA-rated mortgage-backed securities have been downgraded to a CCC rating. As a result, the price of each affected mortgage-backed security has declined 80 percent from its original level. How do we construct* $\bar{\mathbf{A}}$*,* $\mathbf{w}_i$ *for all* $i \in \{1, \ldots, M\}$*,* $\mathbf{w}_{agg}$*, and* $\bar{\mathbf{p}}$*?*

We start with the $M \times N$ matrix $\mathbf{A}$, and we identify the indices $j \in \{1, \ldots, N\}$ of AAA-rated mortgage-backed securities. There are L such unique securities. We consider two mortgage-backed securities to be the same, and therefore not unique, if they happen to be identically formed from the same underlying pool of mortgages; it is therefore possible for a financial institution to hold more than one unit of a particular AAA-rated mortgage-backed security. We construct the $M \times L$ matrix $\bar{\mathbf{A}}$ by extracting the relevant columns of AAA-rated mortgage-backed security portfolio holdings from the matrix $\mathbf{A}$. We set $\mathbf{w}_i = \left([\bar{\mathbf{A}}]_{i*}\right)^T$ as the updated relevant portfolio vector for financial institution i, and we set $\mathbf{w}_{agg} = \bar{\mathbf{A}}^T\mathbf{1}$ as the updated relevant portfolio vector for the overall financial system. Meanwhile, we construct the $L \times 1$ price vector $\bar{\mathbf{p}}$, whose elements are the prices of AAA-rated mortgage-backed securities. Quantity $\mathbf{w}_i^T\bar{\mathbf{p}}$ is the initial value of the AAA-rated mortgage-backed security portfolio for financial institution i, and $\mathbf{w}_{agg}^T\bar{\mathbf{p}}$ is the initial value of the AAA-rated mortgage-backed security portfolio for the entire financial system. We define $\boldsymbol{\epsilon}$ as the $L \times 1$ vector capturing potential changes to the prices of these mortgage-backed securities. We have $[\boldsymbol{\epsilon}]_\ell = -0.80$ for forty percent of all unique AAA-rated mortgage-backed securities and we have $[\boldsymbol{\epsilon}]_\ell = 0$ for sixty percent of all unique AAA-rated mortgage-backed securities. There are $\binom{L}{0.4L}$ such vectors $\boldsymbol{\epsilon}$, for each vector $\boldsymbol{\epsilon}$ identifies the indices of a different subset of AAA-rated mortgage-backed securities getting an 80-percent reduction in price. The value of the

AAA-rated mortgage-backed security portfolio for financial institution i following a specific $\boldsymbol{\epsilon}$-shock is: $\mathbf{w}_i^T (\bar{\mathbf{p}} + \boldsymbol{\epsilon} \circ \bar{\mathbf{p}})$. The value of the AAA-rated mortgage-backed security portfolio for the entire financial system following a specific $\boldsymbol{\epsilon}$-shock is: $\mathbf{w}_{agg}^T (\bar{\mathbf{p}} + \boldsymbol{\epsilon} \circ \bar{\mathbf{p}})$.

Example 3.2 *Suppose that the category of risk is exchange rate risk. Specifically, one foreign currency depreciates by 15 percent relative to the U.S. dollar. The market value of net assets for each financial institution is priced in U.S. dollars. How do we construct* $\bar{\mathbf{A}}$, $\mathbf{w}_i$ *for all* $i \in \{1, \ldots, M\}$, $\mathbf{w}_{agg}$, $\bar{\mathbf{p}}_i$ *for all* $i \in \{1, \ldots, M\}$, *and* $\bar{\mathbf{p}}_{agg}$*?*

Suppose that there are L total foreign currencies, indexed by $j \in \{1, \ldots, L\}$. We therefore construct an $M \times L$ matrix $\bar{\mathbf{A}}$. Element ij of matrix $\bar{\mathbf{A}}$ equals 0 if financial institution i does not hold any assets or liabilities denominated in foreign currency j; otherwise, element ij of matrix $\bar{\mathbf{A}}$ equals 1. We set $\mathbf{w}_i = \left([\bar{\mathbf{A}}]_{i*}\right)^T$ as the updated relevant portfolio vector for financial institution i. Similarly, we introduce $\mathbf{w}_{agg}$ as the updated relevant portfolio vector for the entire financial system. We set $\left[\mathbf{w}_{agg}\right]_j = 0$ if the overall financial system does not hold any assets or liabilities denominated in foreign currency j; otherwise, we set $\left[\mathbf{w}_{agg}\right]_j = 1$. We next construct $\bar{\mathbf{p}}_i$ and $\bar{\mathbf{p}}_{agg}$. Element $[\bar{\mathbf{p}}_i]_j$ is equal to the initial market value, in U.S. dollars, of net assets denominated in foreign currency j for financial institution i. Specifically, define $\mathcal{K} \subseteq \{1, \ldots, N\}$ as the set of indices for securities denominated in foreign currency j. Then, $[\bar{\mathbf{p}}_i]_j = \sum_{k \in \mathcal{K}} [\mathbf{A}]_{ik} [\mathbf{p}]_k$. Quantity $\mathbf{w}_i^T \bar{\mathbf{p}}_i$ is the initial market value, in U.S. dollars, of all net assets denominated in foreign currencies for financial institution i. Meanwhile, element $\left[\bar{\mathbf{p}}_{agg}\right]_j$ is equal to the initial market value, in U.S. dollars, of net assets denominated in foreign currency j for the entire financial system. We therefore set $\left[\bar{\mathbf{p}}_{agg}\right]_j = \sum_{k \in \mathcal{K}} \left[\mathbf{1}^T \mathbf{A}\right]_k [\mathbf{p}]_k$. Quantity $\mathbf{w}_{agg}^T \bar{\mathbf{p}}_{agg}$ is the initial market value, in U.S. dollars, of all net assets denominated in foreign currencies for the entire financial system. We define $\boldsymbol{\epsilon}$ as the $L \times 1$ vector capturing potential changes to the U.S-dollar prices of foreign securities following the exchange rate shock. We set $[\boldsymbol{\epsilon}]_\ell = -0.15$ for the one foreign currency ℓ facing a 15-percent depreciation relative to the U.S. dollar, and for all $r \neq \ell$, we set $[\boldsymbol{\epsilon}]_r = 0$. There are L such vectors $\boldsymbol{\epsilon}$, where each vector $\boldsymbol{\epsilon}$ identifies a different foreign currency depreciating relative to the U.S. dollar. The value, in U.S. dollars, of all net assets denominated in foreign currencies for financial institution i following a specific $\boldsymbol{\epsilon}$-shock is: $\mathbf{w}_i^T \left(\bar{\mathbf{p}}_i + \boldsymbol{\epsilon} \circ \bar{\mathbf{p}}_i\right)$. The value, in U.S. dollars, of all net assets denominated in foreign currencies for the overall financial system following a specific $\boldsymbol{\epsilon}$-shock is: $\mathbf{w}_{agg}^T \left(\bar{\mathbf{p}}_{agg} + \boldsymbol{\epsilon} \circ \bar{\mathbf{p}}_{agg}\right)$.

Example 3.3 *Suppose that the category of risk is industry risk. Specifically, the prices of securities in one industry decline by 20 percent, the prices of securities in another industry decline by 10 percent, and the prices of securities in a third industry increase by 12 percent. Securities have been issued for every industry. How do we construct* $\bar{\mathbf{A}}$, $\mathbf{w}_i$ *for all* $i \in \{1, \ldots, M\}$, $\mathbf{w}_{agg}$, $\bar{\mathbf{p}}_i$ *for all* $i \in \{1, \ldots, M\}$, *and* $\bar{\mathbf{p}}_{agg}$*?*

Suppose that there are L total industries, with each industry indexed by $j \in \{1, \ldots, L\}$. We therefore construct an $M \times L$ matrix $\bar{\mathbf{A}}$. Element ij of matrix $\bar{\mathbf{A}}$ equals 0 if financial institution i does not hold any assets or liabilities from industry j; otherwise, element ij of matrix $\bar{\mathbf{A}}$ equals 1. We set $\mathbf{w}_i = \left([\bar{\mathbf{A}}]_{i*}\right)^T$ as the updated relevant portfolio vector for financial institution i. Now, every security from industry j is held by some financial institution in the financial system, so $\mathbf{w}_{agg} = \mathbf{1}_{L \times 1}$ is the updated relevant portfolio vector for the overall financial system. We next construct $\bar{\mathbf{p}}_i$ and $\bar{\mathbf{p}}_{agg}$. Element $[\bar{\mathbf{p}}_i]_j$ is equal to the initial market value of net assets in industry j for financial institution i. Specifically, define $\mathcal{K} \subseteq \{1, \ldots, N\}$ as the set of indices for securities in industry j. Then, $[\bar{\mathbf{p}}_i]_j = \sum_{k \in \mathcal{K}} [\mathbf{A}]_{ik} [\mathbf{p}]_k$. Quantity

$\mathbf{w}_i^T \bar{\mathbf{p}}_i$ is the initial market value of net assets for financial institution i. Meanwhile, element $\left[\bar{\mathbf{p}}_{agg}\right]_j$ is equal to the initial market value of net assets in industry j for the entire financial system. We therefore set $\left[\bar{\mathbf{p}}_{agg}\right]_j = \sum_{k\in\mathcal{K}} \left[\mathbf{1}^T\mathbf{A}\right]_k [\mathbf{p}]_k$. Quantity $\mathbf{w}_{agg}^T \bar{\mathbf{p}}_{agg}$ is the initial market value of net assets for the entire financial system. We define $\boldsymbol{\epsilon}$ as the $L \times 1$ vector capturing potential changes to the prices of securities for different industries. We set $[\boldsymbol{\epsilon}]_\ell = -0.20$ for one industry ℓ, we set $[\boldsymbol{\epsilon}]_r = -0.10$ for a second industry r, and we set $[\boldsymbol{\epsilon}]_s = 0.12$ for a third industry s. There are $\frac{L!}{(L-3)!}$ such distinct vectors $\boldsymbol{\epsilon}$. The market value of net assets for financial institution i following a specific $\boldsymbol{\epsilon}$-shock is: $\mathbf{w}_i^T \left(\bar{\mathbf{p}}_i + \boldsymbol{\epsilon} \circ \bar{\mathbf{p}}_i\right)$. The market value of net assets for the entire financial system following a specific $\boldsymbol{\epsilon}$-shock is: $\mathbf{w}_{agg}^T \left(\bar{\mathbf{p}}_{agg} + \boldsymbol{\epsilon} \circ \bar{\mathbf{p}}_{agg}\right)$.

The previous three examples consider three different categories of risk. Categories of risk can differ in the manner by which they affect securities prices. I therefore map each category of risk to one of four possible risk environments. These four risk environments collectively represent all of the possible ways that categories of risk can shock securities prices. We now proceed to show, within each risk environment, how to compute the distribution of possible changes to net assets for each individual financial institution and how to compute the distribution of possible changes to net assets for the overall financial system. By developing this set of mathematics for each individual risk environment, we are able to compute distributions of possible changes in net assets for individual financial institutions and the overall financial system for all of the categories of risk that map to that particular risk environment.

3.2.3 First Risk Environment: Absolute Price Shocks, Same Across Securities Clusters

We begin by describing and analyzing the first way that a category of risk can change the market value of net assets. We assume that $\ell \in \{1, \ldots, L\}$ clusters of securities experience a shock δ to their overall value. If each cluster only contains one security, then $\ell \in \{1, \ldots, L\}$ securities are individually experiencing a price shock δ. $\delta < 0$ represents a negative shock to price and/or value, while $\delta > 0$ represents a positive shock.[7] For example, let's suppose that the first ℓ clusters of securities are experiencing a shock δ. Then, the change in net assets for financial institution i is $\mathbf{w}_i^T \boldsymbol{\epsilon}$, and the aggregate change in net assets is $\mathbf{w}_{agg}^T \boldsymbol{\epsilon}$, where $\mathbf{w}_i$ is the $L \times 1$ updated relevant portfolio vector for financial institution i, $\mathbf{w}_{agg}$ is the $L \times 1$ updated relevant portfolio vector for the overall financial system, $[\boldsymbol{\epsilon}]_j = \delta$ for $j \in \{1, \ldots, L\}$ and $[\boldsymbol{\epsilon}]_j = 0$ for $j \in \{\ell + 1, \ldots, L\}$. The vector $\boldsymbol{\epsilon}$ that we just defined is capturing one way that the δ-shocks can manifest themselves. There are many more possible vectors $\boldsymbol{\epsilon} \equiv \boldsymbol{\epsilon}(L, \ell)$ for which ℓ clusters receive a δ-shock and the remaining $L - \ell$ clusters receive a shock of zero. $E(L, \ell)$ is the entire set of such vectors; the cardinality of $E(L, \ell)$ is $\binom{L}{\ell}$. $\boldsymbol{\epsilon}(L, \ell) \in E(L, \ell)$ and $\boldsymbol{\epsilon}'(L, \ell) \in E(L, \ell)$ are distinguished by the indices of the ℓ clusters receiving a δ-shock.

Given this first risk environment, in which ℓ clusters of securities have received a δ-shock, we would like to construct the probability distribution of possible changes in the market value of net assets for financial institution i, $\forall i \in \{1, \ldots, M\}$, and we would like to construct the probability distribution of possible changes in the market value of net assets for the overall financial system. To compute

[7] We assume that, when $\delta < 0$, the magnitude of δ is small enough that securities prices remain non-negative following the δ-shock.

these probability distributions and solve for their statistical features, we introduce additional notation. We define $\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, L, \ell) = \mathbf{w}_i^T \boldsymbol{\epsilon}(L, \ell)$ as the change in net assets for institution i for a given configuration of shocks $\boldsymbol{\epsilon}(L, \ell)$. We define $\pi_{agg}(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell) = \mathbf{w}_{agg}^T \boldsymbol{\epsilon}(L, \ell)$ as the change in net assets for the entire financial system for a given configuration of shocks $\boldsymbol{\epsilon}(L, \ell)$. Random variable $\Pi_i(\mathbf{w}_i, L, \ell)$ has a configuration-specific realization $\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, L, \ell)$, and random variable $\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)$ has a configuration-specific realization $\pi_{agg}(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell)$. We assume that each configuration of shocks among the L clusters of securities is equally likely. Accordingly, corresponding to random variable $\Pi_i(\mathbf{w}_i, L, \ell)$ is the CDF

$$G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t) = \frac{1}{|E(L, \ell)|} \sum_{\boldsymbol{\epsilon}(L,\ell) \in E(L,\ell)} \mathbb{1}_{\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, L, \ell) \leq t}$$

with PMF $g_{\Pi_i(\mathbf{w}_i, L, \ell)}(t)$, and corresponding to random variable $\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)$ is CDF

$$G_{\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}(t) = \frac{1}{|E(L, \ell)|} \sum_{\boldsymbol{\epsilon}(L,\ell) \in E(L,\ell)} \mathbb{1}_{\pi_{agg}(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell) \leq t}$$

with PMF $g_{\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}(t)$. We define random variable W_i with realization $[\mathbf{w}_i]_j$, and we define random variable W_{agg} with realization $[\mathbf{w}_{agg}]_j$. We assume that each realization is equally likely. The elements $[\mathbf{w}_i]_j$ and $[\mathbf{w}_{agg}]_j$ are not themselves random; the elements in the vectors $\mathbf{w}_i$ and $\mathbf{w}_{agg}$ are indeed entirely fixed. We simply introduce random variables W_{agg} and W_i to make certain mathematical expressions more compact. We also set $\mathbf{1}^T \mathbf{w}_i = k_i$, and we set $\mathbf{1}^T \mathbf{w}_{agg} = k_{agg}$.

We can now solve for the statistical features of $\Pi_i(\mathbf{w}_i, L, \ell)$ and $\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)$. The next proposition characterizes their first moments:

Proposition 3.1 *The average change in net assets for financial institution i is:*

$$E\Pi_i(\mathbf{w}_i, L, \ell) = \frac{k_i \ell}{L} \delta,$$

and the average change in net assets for the financial system is:

$$E\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \frac{k_{agg}\ell}{L}\delta.$$

The variances of the distributions capturing possible changes in net assets are as follows:

Proposition 3.2 *The change in net assets for financial institution i has a variance of:*

$$\operatorname{Var}\Pi_i\left(\mathbf{w}_i, L, \ell\right) = \delta^2 \frac{\ell}{L}\left(1 - \frac{\ell}{L}\right)\frac{L}{L-1} L \operatorname{Var} W_i,$$

and the change in net assets for the entire financial system has a variance of:

$$\operatorname{Var}\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \delta^2 \frac{\ell}{L}\left(1 - \frac{\ell}{L}\right)\frac{L}{L-1} L \operatorname{Var} W_{agg}.$$

Specifically, $\operatorname{Var} W_i = \frac{1}{L}\sum_{j=1}^{L}\left(\left[\mathbf{w}_i\right]_j - \frac{k_i}{L}\right)^2$ and $\operatorname{Var} W_{agg} = \frac{1}{L}\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j - \frac{k_{agg}}{L}\right)^2$ are population variances.

We can additionally compute the lower and upper bounds on the supports of these distributions. These bounds tell us the range of possible changes in the market value of net assets for individual financial institutions and the overall financial system given that ℓ clusters experience a δ-shock:

Proposition 3.3 *Construct the ordered multiset* $\left\{\tilde{w}_j\right\}_{j=1}^{L}$ *from the elements of* $\mathbf{w}_i$ *so that* $\tilde{w}_j \leq \tilde{w}_{j'}$ *whenever* $j \leq j'$*. When* $\delta < 0$*, the lower and upper bounds on the distribution of possible changes to net assets for institution i are:*

$$\min \operatorname{supp} \Pi_i\left(\mathbf{w}_i, L, \ell\right) = \delta \sum_{j=L-\ell+1}^{L} \tilde{w}_j \quad \textit{and}$$

$$\max \operatorname{supp} \Pi_i\left(\mathbf{w}_i, L, \ell\right) = \delta \sum_{j=1}^{\ell} \tilde{w}_j.$$

Now construct the ordered multiset $\left\{\tilde{x}_j\right\}_{j=1}^{L}$ *from the elements of* $\mathbf{w}_{agg}$ *so that* $\tilde{x}_j \leq \tilde{x}_{j'}$ *whenever* $j \leq j'$*. When* $\delta < 0$*, the lower and upper bounds on the distribution of possible* *changes to net assets for the financial system are:*

$$\min \operatorname{supp} \Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \delta \sum_{j=L-\ell+1}^{L} \tilde{x}_j \ \text{ and}$$

$$\max \operatorname{supp} \Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \delta \sum_{j=1}^{\ell} \tilde{x}_j.$$

We also want to construct asymptotic expansions that approximate the distribution of possible changes in net assets for financial institution i, $\forall i \in \{1, \ldots, M\}$ and the distribution of possible changes in net assets for the overall financial system. In particular, we are interested in approximating $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$ for all $i \in \{1, \ldots, M\}$ and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$. We first introduce the function $J(\widehat{\mathbf{w}}, L, \ell, t)$:

$$J(\widehat{\mathbf{w}}, L, \ell, t) = \Phi(t) - H_2(t)\phi(t)C_1\sum_{j=1}^{L}\widehat{w}_j^3 - H_3(t)\phi(t)\left[C_2\left(\sum_{j=1}^{L}\widehat{w}_j^4 - \frac{3}{L}\right) - \frac{1}{4L}\right] - H_5(t)\phi(t)C_3\left(\sum_{j=1}^{L}\widehat{w}_j^3\right)^2,$$

where $C_1 = \frac{1-\frac{2\ell}{L}}{6\left(\frac{\ell}{L}\left(1-\frac{\ell}{L}\right)\right)^{1/2}}$, $C_2 = \frac{1-6\left(\frac{\ell}{L}\right)\left(1-\frac{\ell}{L}\right)}{24\left(\frac{\ell}{L}\right)\left(1-\frac{\ell}{L}\right)}$, $C_3 = \frac{\left(1-\frac{2\ell}{L}\right)^2}{72\left(\frac{\ell}{L}\right)\left(1-\frac{\ell}{L}\right)}$, $\phi(t) = \Phi'(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$, and $H_j(t)\phi(t) = (-1)^j \frac{d^j}{dt^j}\phi(t)$. When we are interested in approximating $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$, we set $\widehat{w}_j = \frac{[\mathbf{w}_i]_j - EW_i}{\sqrt{L \operatorname{Var} W_i}}$. When we are instead interested in approximating $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$, we set $\widehat{w}_j = \frac{[\mathbf{w}_{agg}]_j - EW_{agg}}{\sqrt{L \operatorname{Var} W_{agg}}}$.

Proposition 3.4 *Provided that condition (c) holds, for all $i \in \{1, \ldots, M\}$,*

$$\left| G_{\frac{\Pi_i(\mathbf{w}_i,L,\ell) - E\Pi_i(\mathbf{w}_i,L,\ell)}{\left(\operatorname{Var}\Pi_i(\mathbf{w}_i,L,\ell)\right)^{1/2}}}(t) - J(\widehat{\mathbf{w}}, L, \ell, t) \right| < C_4 \times \sum_{j=1}^{L} \left|\widehat{w}_j\right|^5$$

with $\widehat{w}_j = \frac{[\mathbf{w}_i]_j - EW_i}{\sqrt{L \operatorname{Var} W_i}}$, and

$$\left| G_{\frac{\Pi_{agg}(\mathbf{w}_{agg},L,\ell) - E\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}{\left(\operatorname{Var}\Pi_{agg}(\mathbf{w}_{agg},L,\ell)\right)^{1/2}}}(t) - J(\widehat{\mathbf{w}}, L, \ell, t) \right| < C_4 \times \sum_{j=1}^{L} \left|\widehat{w}_j\right|^5$$

with $\widehat{w}_j = \frac{[\mathbf{w}_{agg}]_j - EW_{agg}}{\sqrt{L \operatorname{Var} W_{agg}}}$ *for all* t, *where* C_4 *is only a function of* $\frac{\ell}{L}$.

Condition (c) (Robinson (1978)) *Given* $C' > 0$, *there exist* $\eta > 0$, $C > 0$, *and* $\kappa > 0$ *not depending on* L *such that, for any fixed* t, *the number of indices* j, *for which* $\left|\widehat{w}_j x - t - 2\widehat{r}\pi\right| > \eta$, *for all* $x \in \left(C' \left[\max_i |\widehat{w}_i|\right]^{-1}, C\left[\sum_{i=1}^{L} |\widehat{w}_i|^5\right]^{-1}\right)$ *and all* $\widehat{r} = 0, \pm 1, \pm 2, \ldots$, *is greater than* κL, *for all* L.

Condition (c) requires that the elements of $\widehat{\mathbf{w}}$ not be clustered over too few values. Accordingly, condition (c) requires that the elements of $\mathbf{w}_i$ for each $i \in \{1, \ldots, M\}$ and $\mathbf{w}_{agg}$ not be clustered over too few values. Given Proposition 3.4, we have that

$$G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t) \approx J\left(\widehat{\mathbf{w}}, L, \ell, \frac{t - E\Pi_i(\mathbf{w}_i, L, \ell)}{(\operatorname{Var} \Pi_i(\mathbf{w}_i, L, \ell))^{1/2}}\right) \quad \forall i \in \{1, \ldots, M\}$$

with $\widehat{w}_j = \frac{[\mathbf{w}_i]_j - EW_i}{\sqrt{L \operatorname{Var} W_i}}$, and

$$G_{\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}(t) \approx J\left(\widehat{\mathbf{w}}, L, \ell, \frac{t - E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}{(\operatorname{Var} \Pi_{agg}(\mathbf{w}_{agg}, L, \ell))^{1/2}}\right)$$

with $\widehat{w}_j = \frac{[\mathbf{w}_{agg}]_j - EW_{agg}}{\sqrt{L \operatorname{Var} W_{agg}}}$. For each individual financial institution, note that $\sum_{j=1}^{L} \widehat{w}_j^3 = L^{-1/2} \operatorname{Skew} W_i$ and $\sum_{j=1}^{L} \widehat{w}_j^4 - \frac{3}{L} = L^{-1} \times (\text{Excess Kurtosis } W_i)$. For the overall financial system, note that $\sum_{j=1}^{L} \widehat{w}_j^3 = L^{-1/2} \operatorname{Skew} W_{agg}$ and $\sum_{j=1}^{L} \widehat{w}_j^4 - \frac{3}{L} = L^{-1} \times (\text{Excess Kurtosis } W_{agg})$. We can therefore approximate $G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t)$ and $G_{\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}(t)$ in terms of the population moments of W_i and W_{agg}, respectively. The asymptotic expansion $J(\widehat{\mathbf{w}}, L, \ell, t)$ is to order $1/L$.

When $\ell = 1$, we can solve for $G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t)$, $\forall i \in \{1, \ldots, M\}$, and $G_{\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}(t)$ exactly. Observe that when $\ell = 1$, $\Pi_i(\mathbf{w}_i, L, \ell) = \delta W_i$ and $\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) = \delta W_{agg}$, so

$$G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t) = \Pr[\Pi_i(\mathbf{w}_i, L, \ell) \leq t] = \Pr[\delta W_i \leq t] = G_{W_i}\left(\frac{t}{\delta}\right) \;\; \forall i \in \{1, \ldots, M\},$$

and

$$G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t) = \Pr\left[\Pi_{agg}\left(\mathbf{w}_{agg},L,\ell\right) \leq t\right] = \Pr\left[\delta W_{agg} \leq t\right] = G_{W_{agg}}\left(\frac{t}{\delta}\right).$$

Example 3.4 *Suppose that the category of risk is sovereign risk. Specifically, one country has a writedown of its sovereign debt, which causes the price of each sovereign bond for that country to decrease 10 dollars. We are interested in the possible changes in net assets for each individual financial institution i, $\forall i \in \{1,\ldots,M\}$, and we are interested in the possible changes in net assets for the overall financial system.*

We would like to compute the statistical features of $\Pi_i(\mathbf{w}_i, L, \ell)$, $\forall i \in \{1,\ldots,M\}$, which captures the possible changes in net assets for each individual financial institution i. We would also like to compute the statistical features of $\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right)$, which captures possible changes in net assets for the overall financial system. We are additionally interested in constructing the probability distributions $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$, $\forall i \in \{1,\ldots,M\}$, and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$. We therefore proceed to solve for all of the necessary variables. There are L total countries, indexed by $j \in \{1,\ldots,L\}$, that have issued sovereign bonds. We introduce $\mathbf{w}_i$ as the relevant portfolio vector for financial institution i, and we introduce $\mathbf{w}_{agg}$ as the relevant portfolio vector for the overall financial system. We set $[\mathbf{w}_i]_j$ equal to the number of sovereign bonds from country j held by financial institution i, and we set $\left[\mathbf{w}_{agg}\right]_j$ equal to the total number of sovereign bonds from country j held by all M financial institutions in the financial system. We now construct the $L \times 1$ price vector $\bar{\mathbf{p}}$. $[\bar{\mathbf{p}}]_j$ is equal to the dollar price of a country-j sovereign bond prior to a potential writedown of debt. With one country writing down its debt, $\ell = 1$. The price shock is $\delta = -10$. The vector $\boldsymbol{\epsilon}(L,\ell)$ identifies the one country writing down its debt. We set $[\boldsymbol{\epsilon}(L,\ell)]_j = -10$ if country j is writing down its sovereign debt, and otherwise we set $[\boldsymbol{\epsilon}(L,\ell)]_j = 0$. There are L possible vectors $\boldsymbol{\epsilon}(L,\ell)$ in $E(L,\ell)$, with each vector $\boldsymbol{\epsilon}(L, \ell)$ identifying a different country that is having a writedown of its sovereign debt. For a given configuration of shocks $\boldsymbol{\epsilon}(L, \ell)$, the change in net assets for institution i is $\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, L, \ell) = \mathbf{w}_i^T \boldsymbol{\epsilon}(L, \ell)$, and the change in net assets for the entire financial system is $\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell\right) = \mathbf{w}_{agg}^T \boldsymbol{\epsilon}(L, \ell)$. Since $\ell = 1$, $G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t) = G_{W_i}\left(\frac{t}{\delta}\right)$ exactly, $\forall i \in \{1, \ldots, M\}$, and $G_{\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right)}(t) = G_{W_{agg}}\left(\frac{t}{\delta}\right)$ exactly.

3.2.4 Second Risk Environment: Percentage Price Shocks, Same Across Securities Clusters

In this second environment, any given category of risk affects securities prices in the following manner. We have L total clusters of securities, and the category of risk stresses $\ell \in \{1, \ldots, L\}$ clusters of securities. In particular, stressed clusters of securities experience the same percentage adjustment to their market values, while non-stressed clusters of securities experience zero adjustment to their market values. We denote $\widehat{\delta}$ as that percentage adjustment to a stressed cluster's market value, or equivalently, that cluster's securities prices; $\widehat{\delta} < 0$ means that there is a negative shock to the cluster's securities prices, while $\widehat{\delta} > 0$ means that there is a positive shock to the cluster's securities prices. We introduce the $L \times 1$ vector $\boldsymbol{\epsilon}(L, \ell)$ to capture price shocks to securities clusters. $[\boldsymbol{\epsilon}(L, \ell)]_j = \widehat{\delta}$ if cluster j is stressed, and otherwise $[\boldsymbol{\epsilon}(L, \ell)]_j = 0$; with there being ℓ total stressed clusters, we must have $\mathbf{1}^T \boldsymbol{\epsilon}(L, \ell) = \widehat{\delta}\ell$. We also introduce the $L \times 1$ price vector $\bar{\mathbf{p}}$. If each cluster of securities j contains only one type of security, $\forall j \in \{1, \ldots, L\}$, then we set $[\bar{\mathbf{p}}]_j$ equal to the initial price of the security in that cluster. However, if at least one cluster contains more than one type of security, then for all $j \in \{1, \ldots, L\}$, we set $[\bar{\mathbf{p}}]_j$ equal to the initial market value of the collection of securities in cluster j. For a given configuration of percentage price shocks $\boldsymbol{\epsilon}(L,\ell)$, the change in the market value of net assets for financial institution i is $\pi_i(\mathbf{w}_i,\boldsymbol{\epsilon},L,\ell) = \mathbf{w}_i^T(\boldsymbol{\epsilon}(L,\ell)\circ\bar{\mathbf{p}})$, $\forall i\in\{1,\ldots,M\}$, and the change in the market value of net assets for the overall financial system is $\pi_{agg}(\mathbf{w}_{agg},\boldsymbol{\epsilon},L,\ell) = \mathbf{w}_{agg}^T(\boldsymbol{\epsilon}(L,\ell)\circ\bar{\mathbf{p}})$.

There are many possible configurations of percentage price shocks among the L clusters of securities. Each configuration is distinguished by the particular subset of clusters receiving a percentage price shock. $E(L,\ell)$ is the set of all possible configurations of percentage price shocks given that ℓ clusters experience a $\widehat{\delta}$ percentage-price shock, and $L-\ell$ clusters experience zero price shock. The cardinality of the set $E(L,\ell)$ is combinatorial: $|E(L,\ell)| = \binom{L}{\ell}$. We are interested in characterizing the statistical properties of the random variable $\Pi_i(\mathbf{w}_i,L,\ell)$, $\forall i\in\{1,\ldots,M\}$, whose realizations are $\pi_i(\mathbf{w}_i,\boldsymbol{\epsilon},L,\ell)$, and we are interested in characterizing the statistical properties of the random variable $\Pi_{agg}(\mathbf{w}_{agg},L,\ell)$, whose realizations are $\pi_{agg}(\mathbf{w}_{agg},\boldsymbol{\epsilon},L,\ell)$. $\Pi_i(\mathbf{w}_i,L,\ell)$ and $\Pi_{agg}(\mathbf{w}_{agg},L,\ell)$ respectively represent the possible changes in net assets for financial institution i and the overall financial system given that ℓ clusters of securities experience a $\widehat{\delta}$ percentage price shock. We would also like to approximate the CDFs $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$ and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$.

To construct these probability distributions and solve for their statistical features, we need to rewrite certain expressions of interest. Define the $L\times 1$ vector $\mathbf{b}(L,\ell)$, with $[\mathbf{b}(L,\ell)]_j = 1$ if $[\boldsymbol{\epsilon}(L,\ell)]_j = \widehat{\delta}$ and $[\mathbf{b}(L,\ell)]_j = 0$ if $[\boldsymbol{\epsilon}(L,\ell)]_j = 0$. The vector $\mathbf{b}(L,\ell)$ identifies the indices of those clusters being stressed. Additionally, define the vectors $\mathbf{v}_i$, $\forall i\in\{1,\ldots,M\}$, and $\mathbf{v}_{agg}$:

$$\mathbf{v}_i = \begin{pmatrix}[\mathbf{w}_i]_1 & \frac{[\mathbf{w}_i]_2[\bar{\mathbf{p}}]_2}{[\bar{\mathbf{p}}]_1} & \ldots & \frac{[\mathbf{w}_i]_L[\bar{\mathbf{p}}]_L}{[\bar{\mathbf{p}}]_1}\end{pmatrix}^T, \text{ and}$$

$$\mathbf{v}_{agg} = \begin{pmatrix}[\mathbf{w}_{agg}]_1 & \frac{[\mathbf{w}_{agg}]_2[\bar{\mathbf{p}}]_2}{[\bar{\mathbf{p}}]_1} & \ldots & \frac{[\mathbf{w}_{agg}]_L[\bar{\mathbf{p}}]_L}{[\bar{\mathbf{p}}]_1}\end{pmatrix}^T.$$

We then establish the following lemma:

Lemma 3.1 *Given the configuration $\boldsymbol{\epsilon}(L,\ell)$ of percentage price shocks, the change in the market value of net assets for financial institution i is:*

$$\mathbf{w}_i^T\left(\boldsymbol{\epsilon}(L,\ell)\circ\bar{\mathbf{p}}\right)=\left(\widehat{\delta}\,[\bar{\mathbf{p}}]_1\right)\mathbf{v}_i^T\mathbf{b}(L,\ell),$$

$\forall i\in\{1,\ldots,M\}$, and the change in the market value of net assets for the entire financial system is:

$$\mathbf{w}_{agg}^T\left(\boldsymbol{\epsilon}(L,\ell)\circ\bar{\mathbf{p}}\right)=\left(\widehat{\delta}\,[\bar{\mathbf{p}}]_1\right)\mathbf{v}_{agg}^T\mathbf{b}(L,\ell).$$

We set $k_i=\mathbf{1}^T\mathbf{v}_i$, $\forall i\in\{1,\ldots,M\}$, and we set $k_{agg}=\mathbf{1}^T\mathbf{v}_{agg}$. In addition, we define random variable V_i with realization $[\mathbf{v}_i]_j$, and we define random variable V_{agg} with realization $\left[\mathbf{v}_{agg}\right]_j$. Each realization is equally likely. We can now solve, in closed form, for the statistical features of $\Pi_i(\mathbf{w}_i,L,\ell)$ and $\Pi_{agg}\left(\mathbf{w}_{agg},L,\ell\right)$. We have the following results:

Proposition 3.5 *The average change in net assets for financial institution i is:*

$$E\Pi_i(\mathbf{w}_i,L,\ell)=\frac{k_i\ell}{L}\widehat{\delta}\,[\bar{\mathbf{p}}]_1,$$

and the average change in net assets for the financial system is:

$$E\Pi_{agg}\left(\mathbf{w}_{agg},L,\ell\right)=\frac{k_{agg}\ell}{L}\widehat{\delta}\,[\bar{\mathbf{p}}]_1$$

Proposition 3.6 *The change in net assets for financial institution i has a variance of:*

$$\operatorname{Var}\Pi_i(\mathbf{w}_i,L,\ell)=\left(\widehat{\delta}\,[\bar{\mathbf{p}}]_1\right)^2\frac{\ell}{L}\left(1-\frac{\ell}{L}\right)\frac{L}{L-1}L\operatorname{Var}V_i,$$

and the change in net assets for the entire financial system has a variance of:

$$\operatorname{Var}\Pi_{agg}\left(\mathbf{w}_{agg},L,\ell\right)=\left(\widehat{\delta}\,[\bar{\mathbf{p}}]_1\right)^2\frac{\ell}{L}\left(1-\frac{\ell}{L}\right)\frac{L}{L-1}L\operatorname{Var}V_{agg}.$$

Specifically, $\text{Var}\,V_i = \frac{1}{L}\sum_{j=1}^{L}\left([\mathbf{v}_i]_j - \frac{k_i}{L}\right)^2$ and $\text{Var}\,V_{agg} = \frac{1}{L}\sum_{j=1}^{L}\left(\left[\mathbf{v}_{agg}\right]_j - \frac{k_{agg}}{L}\right)^2$.

Proposition 3.7 *Construct the ordered multiset* $\{\tilde{v}_j\}_{j=1}^{L}$ *from the elements of* $\mathbf{v}_i$ *so that* $\tilde{v}_j \leq \tilde{v}_{j'}$ *whenever* $j \leq j'$. *When* $\widehat{\delta} < 0$, *the lower and upper bounds on the distribution of possible changes to net assets for financial institution i are:*

$$\min \text{supp}\,\Pi_i\left(\mathbf{w}_i, L, \ell\right) = \widehat{\delta}\left[\bar{\mathbf{p}}\right]_1 \sum_{j=L-\ell+1}^{L} \tilde{v}_j \quad \textit{and}$$

$$\max \text{supp}\,\Pi_i\left(\mathbf{w}_i, L, \ell\right) = \widehat{\delta}\left[\bar{\mathbf{p}}\right]_1 \sum_{j=1}^{\ell} \tilde{v}_j.$$

Now construct the ordered multiset $\{\tilde{x}_j\}_{j=1}^{L}$ *from the elements of* $\mathbf{v}_{agg}$ *so that* $\tilde{x}_j \leq \tilde{x}_{j'}$ *whenever* $j \leq j'$. *When* $\widehat{\delta} < 0$, *the lower and upper bounds on the distribution of possible changes to net assets for the overall financial system are:*

$$\min \text{supp}\,\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \widehat{\delta}\left[\bar{\mathbf{p}}\right]_1 \sum_{j=L-\ell+1}^{L} \tilde{x}_j \quad \textit{and}$$

$$\max \text{supp}\,\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \widehat{\delta}\left[\bar{\mathbf{p}}\right]_1 \sum_{j=1}^{\ell} \tilde{x}_j.$$

Proposition 3.8 *Provided that condition (c) holds, for all* $i \in \{1, \ldots, M\}$,

$$\left| G_{\frac{\Pi_i(\mathbf{w}_i, L, \ell) - E\Pi_i(\mathbf{w}_i, L, \ell)}{\left(\text{Var}\,\Pi_i(\mathbf{w}_i, L, \ell)\right)^{1/2}}}(t) - J\left(\widehat{\mathbf{w}}, L, \ell, t\right) \right| < C_4 \times \sum_{j=1}^{L} \left|\widehat{w}_j\right|^5$$

with $\widehat{w}_j = \frac{[\mathbf{v}_i]_j - EV_i}{\sqrt{L\,\text{Var}\,V_i}}$, *and*

$$\left| G_{\frac{\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) - E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}{\left(\text{Var}\,\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)\right)^{1/2}}}(t) - J\left(\widehat{\mathbf{w}}, L, \ell, t\right) \right| < C_4 \times \sum_{j=1}^{L} \left|\widehat{w}_j\right|^5$$

with $\widehat{w}_j = \frac{\left[\mathbf{v}_{agg}\right]_j - EV_{agg}}{\sqrt{L\,\text{Var}\,V_{agg}}}$ *for all t, where* C_4 *is only a function of* $\frac{\ell}{L}$.

Condition (c) requires that the elements of $\mathbf{v}_i$, for all $i \in \{1, \ldots, M\}$, and the elements of $\mathbf{v}_{agg}$ not be clustered over too few values. Given Proposition 3.8, we have that

$$G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t) \approx J\left(\widehat{\mathbf{w}}, L, \ell, \frac{t - E\Pi_i(\mathbf{w}_i, L, \ell)}{(\operatorname{Var}\Pi_i(\mathbf{w}_i, L, \ell))^{1/2}}\right),\ \forall i \in \{1, \ldots, M\},$$

with $\widehat{w}_j = \frac{[\mathbf{v}_i]_j - EV_i}{\sqrt{L \operatorname{Var} V_i}}$, and

$$G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t) \approx J\left(\widehat{\mathbf{w}}, L, \ell, \frac{t - E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)}{(\operatorname{Var}\Pi_{agg}(\mathbf{w}_{agg}, L, \ell))^{1/2}}\right),$$

with $\widehat{w}_j = \frac{[\mathbf{v}_{agg}]_j - EV_{agg}}{\sqrt{L \operatorname{Var} V_{agg}}}$. For each individual financial institution, note that $\sum_{j=1}^{L} \widehat{w}_j^3 = L^{-1/2}\operatorname{Skew} V_i$ and $\sum_{j=1}^{L} \widehat{w}_i^4 - \frac{3}{L} = L^{-1} \times (\text{Excess Kurtosis}\ V_i)$. For the overall financial system, note that $\sum_{j=1}^{L} \widehat{w}_j^3 = L^{-1/2}\operatorname{Skew} V_{agg}$ and $\sum_{j=1}^{L} \widehat{w}_i^4 - \frac{3}{L} = L^{-1} \times (\text{Excess Kurtosis}\ V_{agg})$. We can therefore approximate the CDFs $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$ and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$ respectively in terms of the population moments of V_i and V_{agg}. The asymptotic expansion $J(\widehat{\mathbf{w}}, L, \ell, t)$ is to order $1/L$.

When $\ell = 1$, we can solve for $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$, $\forall i \in \{1, \ldots, M\}$, and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$ exactly. When $\ell = 1$, $\Pi_i(\mathbf{w}_i, L, \ell) = \widehat{\delta}[\bar{\mathbf{p}}]_1 V_i$ and $\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) = \widehat{\delta}[\bar{\mathbf{p}}]_1 V_{agg}$, so $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t) = G_{V_i}\left(\frac{t}{\widehat{\delta}[\bar{\mathbf{p}}]_1}\right)$ and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t) = G_{V_{agg}}\left(\frac{t}{\widehat{\delta}[\bar{\mathbf{p}}]_1}\right)$.

Example 3.5 *Suppose that the category of risk is exchange rate risk. Specifically, one foreign currency depreciates by 15 percent relative to the U.S. dollar. The market value of net assets for each financial institution is priced in U.S. dollars. We are interested in the possible changes in net assets for each individual financial institution i,* $\forall i \in \{1, \ldots, M\}$*, and we are interested in the possible changes in net assets for the overall financial system.*

We would like to compute the statistical features of $\Pi_i(\mathbf{w}_i, L, \ell)$, $\forall i \in \{1, \ldots, M\}$, which captures the possible changes in net assets for each individual financial institution i. We would also like to compute the statistical features of $\Pi_{agg}(\mathbf{w}_{agg}, L, \ell)$, which captures the possible changes in net assets for the overall financial system. We are additionally interested in constructing the probability distributions $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$, $\forall i \in \{1,\dots,M\}$, and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$. We therefore proceed to solve for all of the necessary variables. There are L total foreign currencies, indexed by $j \in \{1,\dots,L\}$. We introduce $\mathbf{w}_i$ as the $L \times 1$ relevant portfolio vector for financial institution i, and we introduce $\mathbf{w}_{agg}$ as the $L \times 1$ relevant portfolio vector for the overall financial system. We also introduce $\bar{\mathbf{p}}_i$ as the $L \times 1$ relevant price vector for financial institution i, and we introduce $\bar{\mathbf{p}}_{agg}$ as the relevant price vector for the overall financial system. We construct $\mathbf{w}_i$, $\forall i \in \{1,\dots,M\}$, $\mathbf{w}_{agg}$, $\bar{\mathbf{p}}_i$, $\forall i \in \{1,\dots,M\}$, and $\bar{\mathbf{p}}_{agg}$ by following Example 3.2. We then construct $\mathbf{v}_i$, $\forall i \in \{1,\dots,M\}$, from the vectors $\mathbf{w}_i$ and $\bar{\mathbf{p}}_i$, and we construct $\mathbf{v}_{agg}$ from the vectors $\mathbf{w}_{agg}$ and $\bar{\mathbf{p}}_{agg}$. With one foreign currency depreciating by 15 percent relative to the U.S. dollar, $\ell = 1$ and $\widehat{\delta} = -0.15$.[8] We construct $\boldsymbol{\epsilon}(L,\ell)$ by setting $[\boldsymbol{\epsilon}(L,\ell)]_j = -0.15$ for one foreign currency j, and we set $[\boldsymbol{\epsilon}(L,\ell)]_r = 0$ for all other $r \neq j$. The set of all possible configurations of exchange rate shocks is $E(L,\ell)$, with $|E(L,\ell)| = L$. Each possible configuration features a different foreign currency experiencing a 15-percent depreciation. Given an exchange rate shock $\boldsymbol{\epsilon}(L,\ell)$, the change in net assets for financial institution i is $\pi_i(\mathbf{w}_i,\boldsymbol{\epsilon},L,\ell) = \left(\widehat{\delta}\,[\bar{\mathbf{p}}_i]_1\right)\mathbf{v}_i^T\mathbf{b}(L,\ell)$, and the change in net

[8]To demonstrate why we set $\widehat{\delta} = -0.15$, see the following proof: Define the exchange rate $\mathcal{E}^{(q)}_{\$/F}$ as the number of U.S. dollars per unit of foreign currency in time period q; time period 0 precedes depreciation and time period 1 immediately follows depreciation. Also define $P^{(q)}_F$ as the period-q price of a foreign security denominated in foreign currency, and define $P^{(q)}_\$$ as the period-q price of a foreign security denominated in U.S. dollars. With

$$\frac{\mathcal{E}^{(1)}_{\$/F} - \mathcal{E}^{(0)}_{\$/F}}{\mathcal{E}^{(0)}_{\$/F}} = -0.15,$$

so that $\mathcal{E}^{(1)}_{\$/F} = \mathcal{E}^{(0)}_{\$/F} \times (1-0.15)$, we then have

$$P^{(1)}_\$ = P^{(0)}_F \times \mathcal{E}^{(1)}_{\$/F} = P^{(0)}_F \times \mathcal{E}^{(0)}_{\$/F} \times (1-0.15) = P^{(0)}_\$ \times (1-0.15) = P^{(0)}_\$ + \widehat{\delta} P^{(0)}_\$, \text{ setting } \widehat{\delta} = -0.15.$$

assets for the entire financial system is $\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell\right) = \left(\widehat{\delta}\left[\bar{\mathbf{p}}_{agg}\right]_1\right)\mathbf{v}_{agg}^T\mathbf{b}\left(L, \ell\right)$. Moreover, since $\ell = 1$, $G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t) = G_{V_i}\left(\frac{t}{\widehat{\delta}[\bar{\mathbf{p}}_i]_1}\right)$ exactly, $\forall i \in \{1, \ldots, M\}$, and $G_{\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right)}(t) = G_{V_{agg}}\left(\frac{t}{\widehat{\delta}\left[\bar{\mathbf{p}}_{agg}\right]_1}\right)$ exactly.

Example 3.6 *Suppose that the category of risk is credit risk. Specifically, forty percent of all unique AAA-rated mortgage-backed securities have been downgraded to a CCC rating. As a result, the price of each affected mortgage-backed security has declined 80 percent from its original level. We are interested in the possible changes in net assets for each individual financial institution i, $\forall i \in \{1, \ldots, M\}$, and we are interested in the possible changes in net assets for the overall financial system.*

We would like to compute the statistical features of $\Pi_i\left(\mathbf{w}_i, L, \ell\right)$, $\forall i \in \{1, \ldots, M\}$, which captures the possible changes in net assets for each individual financial institution i. We would also like to compute the statistical features of $\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right)$, which captures the possible changes in net assets for the overall financial system. We are additionally interested in constructing asymptotic expansions that approximate the probability distributions $G_{\Pi_i(\mathbf{w}_i, L, \ell)}(t)$, $\forall i \in \{1, \ldots, M\}$, and $G_{\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right)}(t)$. We thus proceed to solve for all of the necessary variables. There are L total unique AAA-rated mortgage-backed securities, indexed by $j \in \{1, \ldots, L\}$. We introduce $\mathbf{w}_i$ as the $L \times 1$ relevant portfolio vector for financial institution i, and we introduce $\mathbf{w}_{agg}$ as the $L \times 1$ relevant portfolio vector for the overall financial system. We also introduce $\bar{\mathbf{p}}$ as the $L \times 1$ relevant price vector. We construct $\mathbf{w}_i$, $\forall i \in \{1, \ldots, M\}$, $\mathbf{w}_{agg}$, and $\bar{\mathbf{p}}$ by following Example 3.1. We then construct $\mathbf{v}_i$, $\forall i \in \{1, \ldots, M\}$, from the vectors $\mathbf{w}_i$ and $\bar{\mathbf{p}}$, and we construct $\mathbf{v}_{agg}$ from the vectors $\mathbf{w}_{agg}$ and $\bar{\mathbf{p}}$. We set $\ell = 0.4L$, and we set $\widehat{\delta} = -0.8$. The $L \times 1$ vector $\boldsymbol{\epsilon}\left(L, \ell\right)$ identifies the indices of those AAA-rated mortgage-backed securities experiencing a reduction in price; $\left[\boldsymbol{\epsilon}\left(L, \ell\right)\right]_j = -0.8$ if AAA-rated mortgage-backed security j has been downgraded to a CCC rating, and otherwise $[\epsilon(L,\ell)]_j = 0$. Each vector $\epsilon(L,\ell)$ identifies a different subset of AAA-rated mortgage-backed securities experiencing a percentage price reduction. There are many such vectors $\epsilon(L,\ell)$ in the set $E(L,\ell)$: $|E(L,\ell)| = \binom{L}{0.4L}$. For any given vector $\epsilon(L,\ell)$, the change in net assets for financial institution i is $\pi_i(\mathbf{w}_i,\epsilon,L,\ell) = \left(\widehat{\delta}[\bar{\mathbf{p}}]_1\right)\mathbf{v}_i^T\mathbf{b}(L,\ell)$, and the change in net assets for the entire financial system is $\pi_{agg}(\mathbf{w}_{agg},\epsilon,L,\ell) = \left(\widehat{\delta}[\bar{\mathbf{p}}]_1\right)\mathbf{v}_{agg}^T\mathbf{b}(L,\ell)$. We can approximate both $G_{\Pi_i(\mathbf{w}_i,L,\ell)}(t)$, $\forall i \in \{1,\ldots,M\}$, and $G_{\Pi_{agg}(\mathbf{w}_{agg},L,\ell)}(t)$ via asymptotic expansion by following Proposition 3.8.

3.2.5 Third Risk Environment: Absolute Price Shocks, Different Across Securities Clusters

In this third environment, the category of risk affects securities prices in a manner different from the previous two environments. In the previous two environments, every stressed cluster of securities had the same price adjustment. For the first environment, every stressed cluster had the same magnitude of adjustment to its price or value, and for the second environment, every stressed cluster had the same percentage adjustment to its price or value. In the present environment, different stressed clusters can potentially have different magnitudes of adjustment to their prices or values. With L total clusters of securities, we introduce the $L \times 1$ vector δ that specifies possible price or value adjustments to the various clusters. Specifically, one cluster's price or value must adjust by $[\delta]_1$, a separate cluster's price or value must adjust by $[\delta]_2$, and so on. The vector δ does not identify the actual indices of the clusters receiving such price adjustments. Rather, it is the $L \times 1$ vector $\epsilon(\delta) = \mathbf{P}\delta$, with permutation matrix $\mathbf{P}$, that specifies how each cluster of securities adjusts in price or value. Namely, the price or value of securities in cluster 1 adjusts by $[\boldsymbol{\epsilon}(\delta)]_1$, the price or value of securities in cluster 2 adjusts by $[\boldsymbol{\epsilon}(\delta)]_2$, and the price or value of securities in cluster j adjusts by $[\boldsymbol{\epsilon}(\delta)]_j$. The change in the market value of net assets for financial institution i is then $\mathbf{w}_i^T \boldsymbol{\epsilon}(\delta)$, and the change in the market value of net assets for the entire financial system is then $\mathbf{w}_{agg}^T \boldsymbol{\epsilon}(\delta)$.

In general, there are many possible ways that a category of risk, encapsulated by δ, can manifest itself in securities prices. By rearranging the elements of δ, we keep the category of risk the same, but we change which clusters of securities receive certain shocks to price or value. $E(\delta)$ is the set of all possible configurations of shocks given δ. To generate different configurations of shocks, we rearrange or permute the elements of δ. Configurations $\boldsymbol{\epsilon}(\delta) = \mathbf{P}\delta$ and $\boldsymbol{\epsilon}'(\delta) = \mathbf{P}'\delta$ must have $\mathbf{P} \neq \mathbf{P}'$ in order to be distinct.

As in the previous environments, we define $\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, \delta) = \mathbf{w}_i^T \boldsymbol{\epsilon}(\delta)$ as the change in net assets for institution $i \in \{1, \ldots, M\}$ given the configuration of shocks $\boldsymbol{\epsilon}(\delta) = \mathbf{P}\delta$, and we define $\pi_{agg}(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \delta) = \mathbf{w}_{agg}^T \boldsymbol{\epsilon}(\delta)$ as the change in net assets for the entire financial system given the configuration of shocks $\boldsymbol{\epsilon}(\delta) = \mathbf{P}\delta$. Note that the arguments of $\pi_i(\cdot)$ and $\pi_{agg}(\cdot)$ have changed from the previous two environments due to the different nature of this third risk environment. We are interested in characterizing the statistical properties of random variables $\Pi_i(\mathbf{w}_i, \delta)$, $\forall i \in \{1, \ldots, M\}$, and $\Pi_{agg}(\mathbf{w}_{agg}, \delta)$, which respectively represent the possible changes in net assets for financial institution i given risk category δ and the possible changes in net assets for the overall financial system given risk category δ.

As before, we define random variable W_i with realization $[\mathbf{w}_i]_j$, and we define random variable W_{agg} with realization $[\mathbf{w}_{agg}]_j$. Furthermore, $\mathbf{1}^T \mathbf{w}_i = k_i$ and $\mathbf{1}^T \mathbf{w}_{agg} = k_{agg}$.

Proposition 3.9 *The average change in net assets for financial institution i given δ is:*

$$E\Pi_i\left(\mathbf{w}_i, \delta\right) = \left(\frac{\mathbf{1}^T\delta}{L}\right) k_i,$$

and the average change in net assets for the entire financial system given δ is:

$$E\Pi_{agg}\left(\mathbf{w}_{agg}, \delta\right) = \left(\frac{\mathbf{1}^T\delta}{L}\right) k_{agg}.$$

Given the category of risk δ, we would next like to compute the variance of the distribution of possible changes in net assets for financial institution i, and we would like to compute the variance of the distribution of possible changes in net assets for the overall financial system. To solve in closed form for these second moments we need to develop some additional notation. We introduce the $L \times 1$ random vector $\mathbf{\Delta}$ whose elements are the random variables Δ_j, $j \in \{1, \ldots, L\}$. Each random variable Δ_j, $j \in \{1, \ldots, L\}$, has the underlying CDF

$$G_{\Delta_j}(t) = \frac{1}{L} \sum_{m=1}^{L} \mathbb{1}_{[\delta]_m \leq t};$$

it's the empirical probability distribution formed from the elements of δ. For each random variable Δ_j, we draw a value from the elements of δ. These draws are done without replacement; when drawing a value for Δ_1, there are L scalars to choose from, while there is only one scalar to choose from when drawing a value for Δ_L. This is what makes the random variables $\Delta_1, \ldots, \Delta_L$ identically distributed but not independent. Note that $E\Delta_j = \frac{\mathbf{1}^T\delta}{L}$ and $\operatorname{Var}\Delta_j = \frac{1}{L}\sum_{m=1}^{L}\left([\delta]_m - E\Delta_j\right)^2$. Then observe that $\Pi_i\left(\mathbf{w}_i, \delta\right) = \mathbf{w}_i^T\mathbf{\Delta}$ and $\Pi_{agg}\left(\mathbf{w}_{agg}, \delta\right) = \mathbf{w}_{agg}^T\mathbf{\Delta}$.

Proposition 3.10 *The change in net assets for financial institution i has a variance of:*

$$\operatorname{Var}\Pi_i(\mathbf{w}_i,\delta) = \left(\operatorname{Var}\Delta_j\right)\sum_{m=1}^{L}\left([\mathbf{w}_i]_m\right)^2 + \left(E\left[\Delta_j\Delta_r\right] - \left(E\Delta_j\right)\left(E\Delta_r\right)\right)\left[(L-1)\sum_{m=1}^{L}\left([\mathbf{w}_i]_m\right)^2 - L^2\operatorname{Var}W_i\right],$$

and the change in net assets for the entire financial system has a variance of:

$$\operatorname{Var}\Pi_{agg}(\mathbf{w}_{agg},\delta) = \left(\operatorname{Var}\Delta_j\right)\sum_{m=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_m\right)^2 + \left(E\left[\Delta_j\Delta_r\right] - \left(E\Delta_j\right)\left(E\Delta_r\right)\right)\left[(L-1)\sum_{m=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_m\right)^2 - L^2\operatorname{Var}W_{agg}\right].$$

We need to compute $E\left[\Delta_j\Delta_r\right]$, keeping in mind that sampling from the elements of δ is done without replacement. When the support of Δ_j is small, it is straightforward to compute $E\left[\Delta_j\Delta_r\right]$ by hand.

We proceed to compute the lower and upper bounds on the supports of $\Pi_i(\mathbf{w}_i,\delta)$ and $\Pi_{agg}\left(\mathbf{w}_{agg},\delta\right)$:

Proposition 3.11 *Construct the ordered multiset* $\left\{\tilde{w}_j\right\}_{j=1}^{L}$ *from the elements of* $\mathbf{w}_i$ *so that* $\tilde{w}_j \le \tilde{w}_{j'}$ *whenever* $j \le j'$. *Also construct the ordered multiset* $\left\{\tilde{\delta}_j\right\}_{j=1}^{L}$ *from the elements of* δ *so that* $\tilde{\delta}_j \le \tilde{\delta}_{j'}$ *whenever* $j \le j'$. *Then,*

$$\min\operatorname{supp}\Pi_i(\mathbf{w}_i,\delta) = \tilde{w}_1\tilde{\delta}_L + \tilde{w}_2\tilde{\delta}_{L-1} + \cdots + \tilde{w}_L\tilde{\delta}_1 \text{ and}$$

$$\max\operatorname{supp}\Pi_i(\mathbf{w}_i,\delta) = \tilde{w}_1\tilde{\delta}_1 + \tilde{w}_2\tilde{\delta}_2 + \cdots + \tilde{w}_L\tilde{\delta}_L.$$

Now construct the ordered multiset $\left\{\tilde{x}_j\right\}_{j=1}^{L}$ *from the elements of* $\mathbf{w}_{agg}$ *so that* $\tilde{x}_j \le \tilde{x}_{j'}$ *whenever* $j \le j'$. *Then,*

$$\min\operatorname{supp}\Pi_{agg}\left(\mathbf{w}_{agg},\delta\right) = \tilde{x}_1\tilde{\delta}_L + \tilde{x}_2\tilde{\delta}_{L-1} + \cdots + \tilde{x}_L\tilde{\delta}_1 \text{ and}$$

$$\max\operatorname{supp}\Pi_{agg}\left(\mathbf{w}_{agg},\delta\right) = \tilde{x}_1\tilde{\delta}_1 + \tilde{x}_2\tilde{\delta}_2 + \cdots + \tilde{x}_L\tilde{\delta}_L.$$

While we cannot compute $G_{\Pi_i(\mathbf{w}_i,\delta)}(t)$ and $G_{\Pi_{agg}(\mathbf{w}_{agg},\delta)}(t)$ exactly by hand, we can instead simulate these probability distributions by randomly drawing vectors $\boldsymbol{\epsilon}(\delta)$ from the set $E(\delta)$ and computing the corresponding quantities $\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, \delta)$ and $\pi_{agg}(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \delta)$.

Example 3.7 *Suppose that the category of risk is solvency risk. Specifically, 100 public companies have filed for Chapter 11 bankruptcy.*[9] *Each of these public companies undergoes corporate debt restructuring. Each public company renegotiates its debt obligations so that a certain amount of debt is forgiven. As a result, the price of each company's bonds decreases by a fixed amount, with the magnitudes of corporate bond price reductions potentially varying across bankrupt firms. Given this form of solvency risk, we are interested in the possible changes in net assets for each individual financial institution* i, $\forall i \in \{1, \ldots, M\}$, *and we are interested in the possible changes in net assets for the overall financial system.*

We would like to compute the statistical features of $\Pi_i(\mathbf{w}_i, \delta)$, $\forall i \in \{1, \ldots, M\}$, which captures possible changes in net assets for each individual financial institution i, and we would like to compute the statistical features of $\Pi_{agg}(\mathbf{w}_{agg}, \delta)$, which captures possible changes in net assets for the overall financial system. We thus proceed to solve for all of the necessary variables. There are L total public companies with issued corporate debt. Each public company is indexed by $j \in \{1, \ldots, L\}$. We introduce $\mathbf{w}_i$ as the $L \times 1$ relevant portfolio vector for financial institution i, and we introduce $\mathbf{w}_{agg}$ as the $L \times 1$ relevant portfolio vector for the overall financial system. We set $[\mathbf{w}_i]_j$ equal to the number of bonds from public company j held by financial institution i, and we set $[\mathbf{w}_{agg}]_j$ equal to the total number of bonds from public company j held by all M financial institutions in the financial system. We next construct the $L \times 1$ price vector $\bar{\mathbf{p}}$. $[\bar{\mathbf{p}}]_j$ is equal to the initial

[9] In 2008, 136 public companies filed for bankruptcy protection. Source: Jones Day, "The Year in Bankruptcy: 2008." Accessed March 27, 2019.

price of a corporate bond issued by public company j. The $L \times 1$ vector δ captures the changes in the dollar prices of corporate bonds for bankrupt public companies. One public company's corporate bond price reduces by $[\delta]_1$, another public company's corporate bond price reduces by $[\delta]_2$, and so on. Since each bankrupt public company experiences a reduction in its corporate bond price, we have $[\delta]_j < 0$ for $j \in \{1, \ldots, 100\}$. Meanwhile, there is no change in corporate bond prices for the remaining public companies, so $[\delta]_j = 0$ for $j \in \{101, \ldots, L\}$. There are many possible configurations of price shocks among public companies. The $L \times 1$ vector $\boldsymbol{\epsilon}(\delta) = \mathbf{P}\delta$, with permutation matrix $\mathbf{P}$, maps the set of price shocks encapsulated in the vector δ to the set of public companies. Public company 1 experiences a corporate debt price shock of $[\boldsymbol{\epsilon}(\delta)]_1$, public company 2 experiences a corporate debt price shock of $[\boldsymbol{\epsilon}(\delta)]_2$, and public company j experiences a corporate debt price shock of $[\boldsymbol{\epsilon}(\delta)]_j$. For a given configuration of price shocks $\boldsymbol{\epsilon}(\delta)$, the change in net assets for financial institution i is $\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, \delta) = \mathbf{w}_i^T \boldsymbol{\epsilon}(\delta)$, and the change in net assets for the entire financial system is $\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \delta\right) = \mathbf{w}_{agg}^T \boldsymbol{\epsilon}(\delta)$. The set of all possible configurations of price shocks is $E(\delta)$. If each restructured company has a different corporate bond price reduction, then there are $|E(\delta)| = \frac{L!}{(L-100)!}$ unique configurations of price shocks among all L public companies with issued corporate debt.

3.2.6 Fourth Risk Environment: Percentage Price Shocks, Different Across Securities Clusters

In this final environment, any given category of risk affects securities prices in the following manner. Stressed clusters of securities face percentage adjustments to their market values. Unlike the second environment, different stressed clusters of securities can face different percentage adjustments. With L total clusters of securities, we introduce the $L \times 1$ price vector $\bar{\mathbf{p}}$. If each cluster of securities j contains only one type of security, $\forall j \in \{1, \ldots, L\}$, then we set $[\bar{\mathbf{p}}]_j$ equal to the initial price of the security in that cluster. Meanwhile, if at least one cluster contains more than one type of security, then for all $j \in \{1, \ldots, L\}$, we set $[\bar{\mathbf{p}}]_j$ equal to the initial market value of the collection of securities in cluster j. We also introduce the $L \times 1$ vector $\widehat{\delta}$, which specifies possible percentage adjustments to the values of various clusters. Specifically, one cluster's value experiences a percentage adjustment $\left[\widehat{\delta}\right]_1$, another cluster's value experiences a percentage adjustment $\left[\widehat{\delta}\right]_2$, and so on. Equivalently, the prices of securities in one cluster each experience a percentage adjustment $\left[\widehat{\delta}\right]_1$, the prices of securities in another cluster each experience a percentage adjustment $\left[\widehat{\delta}\right]_2$, and so on. The vector $\widehat{\delta}$ does not identify the actual indices of the clusters receiving such price adjustments. Rather, it is the $L \times 1$ vector $\boldsymbol{\epsilon}\left(\widehat{\delta}\right) = \mathbf{P}\widehat{\delta}$, with permutation matrix $\mathbf{P}$, that specifies how each cluster of securities adjusts in value. Namely, securities in cluster 1 experience a percentage adjustment $\left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_1$ to their value, securities in cluster 2 experience a percentage adjustment $\left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_2$ to their value, and securities in cluster j experience a percentage adjustment $\left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_j$ to their value. Given this configuration of shocks $\boldsymbol{\epsilon}\left(\widehat{\delta}\right)$, the change in the market value of net assets for financial institution i is $\mathbf{w}_i^T\left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right)$, and the change in the market value of net assets for the overall financial system is $\mathbf{w}_{agg}^T\left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right)$.

There are many possible ways that a category of risk, encapsulated by $\widehat{\delta}$, can manifest itself in securities prices. By rearranging the elements of $\widehat{\delta}$, we keep the category of risk the same, but we change which clusters of securities receive certain percentage shocks to price or value. $E\left(\widehat{\delta}\right)$ is the set of all possible configurations of shocks given $\widehat{\delta}$. To generate different configurations of shocks, we rearrange or permute the elements of $\widehat{\delta}$. Configurations $\boldsymbol{\epsilon}\left(\widehat{\delta}\right) = \mathbf{P}\widehat{\delta}$ and $\boldsymbol{\epsilon}'\left(\widehat{\delta}\right) = \mathbf{P}'\widehat{\delta}$ must have $\mathbf{P} \neq \mathbf{P}'$ in order to be distinct. We define $\pi_i\left(\mathbf{w}_i, \boldsymbol{\epsilon}, \widehat{\delta}\right) = \mathbf{w}_i^T\left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right)$ as the change in net assets for financial institution $i \in \{1, \ldots, M\}$ given the configuration of shocks $\boldsymbol{\epsilon}\left(\widehat{\delta}\right) = \mathbf{P}\widehat{\delta}$, and we define $\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \widehat{\delta}\right) = \mathbf{w}_{agg}^T\left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right)$ as the change in net assets for the overall financial system given the configuration of shocks $\boldsymbol{\epsilon}\left(\widehat{\delta}\right) = \mathbf{P}\widehat{\delta}$. We are interested in characterizing the statistical properties of the random variable $\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right)$, $\forall i \in \{1, \ldots, M\}$, whose realizations are $\pi_i\left(\mathbf{w}_i, \boldsymbol{\epsilon}, \widehat{\delta}\right)$, and we are interested in characterizing the statistical properties of the random variable $\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right)$, whose realizations are $\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \widehat{\delta}\right)$. $\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right)$ represents the possible changes in net assets for financial institution i given risk category $\widehat{\delta}$, and $\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right)$ represents the possible changes in net assets for the overall financial system given risk category $\widehat{\delta}$.

To solve for the statistical features of $\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right)$, $\forall i \in \{1, \ldots, M\}$, and $\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right)$, we first need to introduce the additional variables $\mathbf{v}_i$, $\forall i \in \{1, \ldots, M\}$, and $\mathbf{v}_{agg}$, which we define below:

$$\mathbf{v}_i = \begin{pmatrix} [\mathbf{w}_i]_1 & \frac{[\mathbf{w}_i]_2[\bar{\mathbf{p}}]_2}{[\bar{\mathbf{p}}]_1} & \ldots & \frac{[\mathbf{w}_i]_L[\bar{\mathbf{p}}]_L}{[\bar{\mathbf{p}}]_1} \end{pmatrix}^T, \text{ and}$$

$$\mathbf{v}_{agg} = \begin{pmatrix} \left[\mathbf{w}_{agg}\right]_1 & \frac{\left[\mathbf{w}_{agg}\right]_2[\bar{\mathbf{p}}]_2}{[\bar{\mathbf{p}}]_1} & \ldots & \frac{\left[\mathbf{w}_{agg}\right]_L[\bar{\mathbf{p}}]_L}{[\bar{\mathbf{p}}]_1} \end{pmatrix}^T.$$

We then have the following lemma:

Lemma 3.2 *Given the configuration of shocks $\boldsymbol{\epsilon}\left(\widehat{\delta}\right)$, the change in the market value of net assets for financial institution i is:*

$$\mathbf{w}_i^T\left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right) = [\bar{\mathbf{p}}]_1\, \mathbf{v}_i^T \boldsymbol{\epsilon}\left(\widehat{\delta}\right),$$

$\forall i \in \{1, \ldots, M\}$, and the change in the market value of net assets for the entire financial system is:

$$\mathbf{w}_{agg}^T\left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right) = [\bar{\mathbf{p}}]_1\, \mathbf{v}_{agg}^T \boldsymbol{\epsilon}\left(\widehat{\delta}\right).$$

We let $\mathbf{1}^T\mathbf{v}_i = k_i$, $\forall i \in \{1,\ldots,M\}$, and we let $\mathbf{1}^T\mathbf{v}_{agg} = k_{agg}$. We also define random variable V_i with realization $[\mathbf{v}_i]_j$, and we define random variable V_{agg} with realization $\left[\mathbf{v}_{agg}\right]_j$. Each realization is equally likely, and that allows us to define population moments for V_i and V_{agg}.

The first moments of $\Pi_i\left(\mathbf{w}_i,\widehat{\delta}\right)$ and $\Pi_{agg}\left(\mathbf{w}_{agg},\widehat{\delta}\right)$ are as follows:

Proposition 3.12 *The average change in net assets for financial institution i given $\widehat{\delta}$ is:*

$$E\Pi_i\left(\mathbf{w}_i,\widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\left(\frac{\mathbf{1}^T\widehat{\delta}}{L}\right)k_i,$$

and the average change in net assets for the financial system given $\widehat{\delta}$ is:

$$E\Pi_{agg}\left(\mathbf{w}_{agg},\widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\left(\frac{\mathbf{1}^T\widehat{\delta}}{L}\right)k_{agg}.$$

Given $\widehat{\delta}$, we would next like to compute the variance of the distribution of possible changes in net assets for financial institution i, and we would like to compute the variance of the distribution of possible changes in net assets for the overall financial system. To solve in closed form for these second moments, we need to develop some additional notation. We introduce the $L \times 1$ random vector $\widehat{\mathbf{\Delta}}$ whose elements are the random variables $\widehat{\Delta}_j$, $j \in \{1,\ldots,L\}$. Each random variable $\widehat{\Delta}_j$, $j \in \{1,\ldots,L\}$ has the underlying CDF

$$G_{\widehat{\Delta}_j}(t) = \frac{1}{L}\sum_{m=1}^{L}\mathbb{1}_{[\widehat{\delta}]_m \le t};$$

it's the empirical probability distribution formed from the elements of $\widehat{\delta}$. For each random variable $\widehat{\Delta}_j$, we draw a value from the elements of $\widehat{\delta}$. These draws are done without replacement; when drawing a value for $\widehat{\Delta}_1$, there are L scalars to choose from, while there is only one scalar to choose from when drawing a value for $\widehat{\Delta}_L$. This is what makes the random variables $\widehat{\Delta}_1,\ldots,\widehat{\Delta}_L$ identically distributed but not independent. The first two population moments for $\widehat{\Delta}_j$ are $E\widehat{\Delta}_j = \frac{\mathbf{1}^T\widehat{\delta}}{L}$ and $\operatorname{Var}\widehat{\Delta}_j = \frac{1}{L}\sum_{m=1}^{L}\left(\left[\widehat{\delta}\right]_m - E\widehat{\Delta}_j\right)^2$. Note that $\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\, \mathbf{v}_i^T\widehat{\boldsymbol{\Delta}}$ and $\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\, \mathbf{v}_{agg}^T\widehat{\boldsymbol{\Delta}}$.

Proposition 3.13 *Given $\widehat{\delta}$, the change in net assets for financial institution i has a variance of:*

$$\begin{aligned}\operatorname{Var}\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right) = ([\bar{\mathbf{p}}]_1)^2 \times \Bigg(&\left(\operatorname{Var}\widehat{\Delta}_j\right)\sum_{m=1}^{L}([\mathbf{v}_i]_m)^2 \\ &+ \left(E\left[\widehat{\Delta}_j\widehat{\Delta}_r\right] - \left(E\widehat{\Delta}_j\right)\left(E\widehat{\Delta}_r\right)\right)\left[(L-1)\sum_{m=1}^{L}([\mathbf{v}_i]_m)^2 - L^2\operatorname{Var}V_i\right]\Bigg),\end{aligned}$$

and the change in net assets for the entire financial system has a variance of:

$$\begin{aligned}\operatorname{Var}\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right) = ([\bar{\mathbf{p}}]_1)^2 \times \Bigg(&\left(\operatorname{Var}\widehat{\Delta}_j\right)\sum_{m=1}^{L}\left(\left[\mathbf{v}_{agg}\right]_m\right)^2 \\ &+ \left(E\left[\widehat{\Delta}_j\widehat{\Delta}_r\right] - \left(E\widehat{\Delta}_j\right)\left(E\widehat{\Delta}_r\right)\right)\left[(L-1)\sum_{m=1}^{L}\left(\left[\mathbf{v}_{agg}\right]_m\right)^2 - L^2\operatorname{Var}V_{agg}\right]\Bigg),\end{aligned}$$

We must compute $E\left[\widehat{\Delta}_j\widehat{\Delta}_r\right]$, keeping in mind that sampling from the elements of $\widehat{\delta}$ is done without replacement. It is straightforward to compute $E\left[\widehat{\Delta}_j\widehat{\Delta}_r\right]$, especially when $\widehat{\Delta}_j$ has a small support.

We next compute the lower and upper bounds on the supports of $\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right)$, and $\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right)$. These lower and upper bounds determine the range of possible adjustments to net income for each individual financial institution and the overall financial system given $\widehat{\delta}$.

Proposition 3.14 *Construct the ordered multiset $\left\{\tilde{v}_j\right\}_{j=1}^{L}$ from the elements of $\mathbf{v}_i$ so that $\tilde{v}_j \leq \tilde{v}_{j'}$ whenever $j \leq j'$. Also construct the ordered multiset $\left\{\tilde{\delta}_j\right\}_{j=1}^{L}$ from the elements of $\widehat{\delta}$ so that $\tilde{\delta}_j \leq \tilde{\delta}_{j'}$ whenever $j \leq j'$. Then*

$$\min\operatorname{supp}\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1 \times \left(\tilde{v}_1\tilde{\delta}_L + \tilde{v}_2\tilde{\delta}_{L-1} + \cdots + \tilde{v}_L\tilde{\delta}_1\right) \text{ and}$$

$$\max \text{supp}\, \Pi_i \left(\mathbf{w}_i, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1 \times \left(\tilde{v}_1\tilde{\delta}_1 + \tilde{v}_2\tilde{\delta}_2 + \cdots + \tilde{v}_L\tilde{\delta}_L\right).$$

Now construct the ordered multiset $\{\tilde{x}_j\}_{j=1}^{L}$ *from the elements of* $\mathbf{v}_{agg}$ *so that* $\tilde{x}_j \leq \tilde{x}_{j'}$ *whenever* $j \leq j'$*. Then,*

$$\min \text{supp}\, \Pi_{agg} \left(\mathbf{w}_{agg}, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1 \times \left(\tilde{x}_1\tilde{\delta}_L + \tilde{x}_2\tilde{\delta}_{L-1} + \cdots + \tilde{x}_L\tilde{\delta}_1\right) \text{ and}$$

$$\max \text{supp}\, \Pi_{agg} \left(\mathbf{w}_{agg}, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1 \times \left(\tilde{x}_1\tilde{\delta}_1 + \tilde{x}_2\tilde{\delta}_2 + \cdots + \tilde{x}_L\tilde{\delta}_L\right).$$

While we cannot compute $G_{\Pi_i\left(\mathbf{w}_i,\widehat{\delta}\right)}(t)$ and $G_{\Pi_{agg}\left(\mathbf{w}_{agg},\widehat{\delta}\right)}(t)$ exactly by hand, we can instead simulate these probability distributions by randomly drawing vectors $\boldsymbol{\epsilon}\left(\widehat{\delta}\right)$ from the set $E\left(\widehat{\delta}\right)$ and computing the corresponding quantities $\pi_i\left(\mathbf{w}_i, \boldsymbol{\epsilon}, \widehat{\delta}\right)$ and $\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \widehat{\delta}\right)$.

Example 3.8 *Suppose that the category of risk is industry risk. Specifically, the prices of securities in one industry decline by 20 percent, the prices of securities in another industry decline by 10 percent, and the prices of securities in a third industry increase by 12 percent. Securities have been issued for every industry. We are interested in the possible changes in net assets for each individual financial institution i,* $\forall i \in \{1, \ldots, M\}$*, and we are interested in the possible changes in net assets for the overall financial system.*

We would like to compute the statistical features of $\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right)$, $\forall i \in \{1, \ldots, M\}$, which captures possible changes in net assets for each individual financial institution i, and we would like to compute the statistical features of $\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right)$, which captures possible changes in net assets for the overall financial system. We therefore solve for all of the necessary variables. There are L total industries, with each industry indexed by $j \in \{1, \ldots, L\}$. We introduce $\mathbf{w}_i$ as the $L \times 1$ relevant portfolio vector for financial institution i, and we introduce $\mathbf{w}_{agg}$ as the $L \times 1$ relevant portfolio vector for the overall financial system. We also introduce $\bar{\mathbf{p}}_i$ as the $L \times 1$ relevant price vector for financial institution i, and we introduce $\bar{\mathbf{p}}_{agg}$ as the relevant price vector for the overall financial system. We construct $\mathbf{w}_i$, $\forall i \in \{1, \ldots, M\}$, $\mathbf{w}_{agg}$, $\bar{\mathbf{p}}_i$, $\forall i \in \{1, \ldots, M\}$, and $\bar{\mathbf{p}}_{agg}$ by following Example 3.3. We then construct $\mathbf{v}_i$, $\forall i \in \{1, \ldots, M\}$, from the vectors $\mathbf{w}_i$ and $\bar{\mathbf{p}}_i$, and we construct $\mathbf{v}_{agg}$ from the vectors $\mathbf{w}_{agg}$ and $\bar{\mathbf{p}}_{agg}$. The $L \times 1$ vector $\widehat{\delta}$ captures percentage changes to the values of securities in each industry. For the category of risk in this example, we set $\left[\widehat{\delta}\right]_1 = -0.20$, $\left[\widehat{\delta}\right]_2 = -0.10$, $\left[\widehat{\delta}\right]_3 = 0.12$, and $\left[\widehat{\delta}\right]_j = 0$ for $j \in \{4, \ldots, L\}$. There are many possible configurations of price shocks among industries. The $L \times 1$ vector $\epsilon\left(\widehat{\delta}\right) = \mathbf{P}\widehat{\delta}$, with permutation matrix $\mathbf{P}$, maps the set of price shocks encapsulated in the vector $\widehat{\delta}$ to the set of industries. The securities from industry 1 experience a percentage price shock of $\left[\epsilon\left(\widehat{\delta}\right)\right]_1$, the securities from industry 2 experience a percentage price shock of $\left[\epsilon\left(\widehat{\delta}\right)\right]_2$, and the securities from industry j experience a percentage price shock of $\left[\epsilon\left(\widehat{\delta}\right)\right]_j$. Given a particular configuration $\epsilon\left(\widehat{\delta}\right)$ of percentage price shocks, the change in net assets for financial institution i is $\pi_i\left(\mathbf{w}_i, \epsilon, \widehat{\delta}\right) = [\bar{\mathbf{p}}_i]_1 \mathbf{v}_i^T \epsilon\left(\widehat{\delta}\right)$, and the change in net assets for the overall financial system is $\pi_{agg}\left(\mathbf{w}_{agg}, \epsilon, \widehat{\delta}\right) = [\bar{\mathbf{p}}_{agg}]_1 \mathbf{v}_{agg}^T \epsilon\left(\widehat{\delta}\right)$. The set of all possible configurations of percentage price shocks is $E\left(\widehat{\delta}\right)$, with there being $\left|E\left(\widehat{\delta}\right)\right| = \frac{L!}{(L-3)!} = L \times (L-1) \times (L-2)$ such unique configurations.

3.2.7 Combining Categories of Risk to Generate Entire Classes of Stress Tests

Thus far, we have been studying four different environments that specify how categories of risk can affect securities prices. Each environment represents a different way that a category of risk can alter net assets for individual financial institutions and the overall financial system. Categories of risk in the first environment generate an absolute adjustment to the values of a certain number of securities clusters, with stressed clusters facing the same magnitude of adjustment. Categories of risk in the second environment generate a percentage adjustment to the values of a certain number of securities clusters, with stressed clusters facing the same percentage adjustment. Categories of risk in the third environment generate potentially different levels of adjustment to the values of securities clusters, and categories of risk in the fourth environment generate potentially different percentage adjustments to the values of securities clusters. Any category of risk fits into one of these four environments. We can map the category of risk to a probability distribution capturing possible balance sheet effects for each individual financial institution, $\forall i \in \{1, \ldots, M\}$, and we can map the category of risk to a probability distribution capturing possible balance sheet effects for the overall financial system. When the individual category of risk fits into the first or second environments, we can either construct the corresponding probability distributions in closed form, or we can construct asymptotic expansions that strongly approximate the CDFs of these probability distributions. When the individual category of risk fits into the third or fourth environments, we can solve in closed form for the major statistical features of the corresponding probability distributions.

Now that we have thoroughly studied individual categories of risk, we would like to construct classes of stress tests. Each class of stress tests features Q categories of risk. We therefore index each category of risk by $q \in \{1, \ldots, Q\}$. We define random variable $\Pi_i^q(\cdot)$ as the change in net assets for financial institution i, $\forall i \in \{1, \ldots, M\}$, given category of risk q, and we define random variable $\Pi_{agg}^q(\cdot)$ as the change in net assets for the overall financial system given category of risk q. Corresponding to these random variables are the CDFs $G_{\Pi_i^q(\cdot)}(t)$ and $G_{\Pi_{agg}^q(\cdot)}(t)$ and the PMFs $g_{\Pi_i^q(\cdot)}(t)$ and $g_{\Pi_{agg}^q(\cdot)}(t)$.

For a given class of stress tests, the random variables Π_i^{class}, $\forall i \in \{1, \ldots, M\}$, and Π_{agg}^{class} respectively capture the entire change in net assets for financial institution i and the entire change in net assets for the overall financial system. We then have the following relationships:

$$\Pi_i^{class} = \Pi_i^1(\cdot) + \Pi_i^2(\cdot) + \cdots + \Pi_i^Q(\cdot), \text{ with}$$

$$g_{\Pi_i^{class}}(t) = \left(g_{\Pi_i^1(\cdot)} * g_{\Pi_i^2(\cdot)} * \cdots * g_{\Pi_i^Q(\cdot)}\right)(t),$$

$\forall i \{1, \ldots, M\}$, and

$$\Pi_{agg}^{class} = \Pi_{agg}^1(\cdot) + \Pi_{agg}^2(\cdot) + \cdots + \Pi_{agg}^Q(\cdot), \text{ with}$$

$$g_{\Pi_{agg}^{class}}(t) = \left(g_{\Pi_{agg}^1(\cdot)} * g_{\Pi_{agg}^2(\cdot)} * \cdots * g_{\Pi_{agg}^Q(\cdot)}\right)(t).$$ [10]

Let $\left|E^{class}\right|$ be the unique number of stress tests in the entire class given that each category of risk q has $|E^q|$ possible configurations of price shocks. Then,

$$\left|E^{class}\right| = \left|E^1\right| \times \left|E^2\right| \times \cdots \times \left|E^Q\right|,$$

which can be extremely large.

We can solve for the major statistical features of Π_i^{class}, $\forall i \in \{1, \ldots, M\}$, and Π_{agg}^{class} in closed form. For all financial institutions $i \in \{1, \ldots, M\}$,

$$E\Pi_i^{class} = \sum_{q=1}^{Q} E\Pi_i^q(\cdot), \quad \text{Var}\,\Pi_i^{class} = \sum_{q=1}^{Q} \text{Var}\,\Pi_i^q(\cdot),$$

$$\text{min supp}\,\Pi_i^{class} = \sum_{q=1}^{Q} \text{min supp}\,\Pi_i^q(\cdot), \text{ and}$$

[10] If more than one category of risk fits into the third environment and/or the fourth environment, we assume that each category of risk causes a percentage change to the *original* value of the securities cluster. We do not consider the case in which percentage changes to the value of a securities cluster are applied sequentially.

$$\max \operatorname{supp} \Pi_i^{class} = \sum_{q=1}^{Q} \max \operatorname{supp} \Pi_i^q (\cdot) .$$

Meanwhile, for Π_{agg}^{class},

$$E\Pi_{agg}^{class} = \sum_{q=1}^{Q} E\Pi_{agg}^q (\cdot) , \quad \operatorname{Var} \Pi_{agg}^{class} = \sum_{q=1}^{Q} \operatorname{Var} \Pi_{agg}^q (\cdot) ,$$

$$\min \operatorname{supp} \Pi_{agg}^{class} = \sum_{q=1}^{Q} \min \operatorname{supp} \Pi_{agg}^q (\cdot) , \text{ and}$$

$$\max \operatorname{supp} \Pi_{agg}^{class} = \sum_{q=1}^{Q} \max \operatorname{supp} \Pi_{agg}^q (\cdot) .$$

When category of risk q fits into the first or second environment, we can approximate $g_{\Pi_i^q(\cdot)} (t)$ and $g_{\Pi_{agg}^q(\cdot)} (t)$ using the relevant asymptotic expansion. If the category of risk fits into the first environment, we establish the following approximations:

$$G_{\Pi_i^q(\cdot)} (t) \approx J \left(\widehat{\mathbf{w}}, L, \ell, \frac{t - E\Pi_i^q (\cdot)}{\left(\operatorname{Var} \Pi_i^q (\cdot)\right)^{1/2}} \right) , \forall i \in \{1, \ldots , M\} ,$$

with $\widehat{w}_j = \frac{[\mathbf{w}_i]_j - EW_i}{\sqrt{L \operatorname{Var} W_i}}$, and

$$G_{\Pi_{agg}^q(\cdot)} (t) \approx J \left(\widehat{\mathbf{w}}, L, \ell, \frac{t - E\Pi_{agg}^q (\cdot)}{\left(\operatorname{Var} \Pi_{agg}^q (\cdot)\right)^{1/2}} \right) ,$$

with $\widehat{w}_j = \frac{[\mathbf{w}_{agg}]_j - EW_{agg}}{\sqrt{L \operatorname{Var} W_{agg}}}$. Depending on the particular category of risk, we need to determine the relevant support for the probability mass functions. If we suppose that the supports of $g_{\Pi_i^q(\cdot)} (t)$, $\forall i \in \{1, \ldots , M\}$, and $g_{\Pi_{agg}^q(\cdot)} (t)$ take integer values, that is, $t \in \mathbb{Z}$, then

$$g_{\Pi_i^q(\cdot)} (t) \approx \lim_{\kappa \uparrow 0.5} J \left(\widehat{\mathbf{w}}, L, \ell, \frac{(t + \kappa) - E\Pi_i^q (\cdot)}{\left(\operatorname{Var} \Pi_i^q (\cdot)\right)^{1/2}} \right) - J \left(\widehat{\mathbf{w}}, L, \ell, \frac{(t - 0.5) - E\Pi_i^q (\cdot)}{\left(\operatorname{Var} \Pi_i^q (\cdot)\right)^{1/2}} \right) ,$$

$\forall i \in \{1, \ldots, M\}$, with $\widehat{w}_j = \frac{[\mathbf{w}_i]_j - EW_i}{\sqrt{L \operatorname{Var} W_i}}$, and

$$g_{\Pi^q_{agg}(\cdot)}(t) \approx \lim_{\kappa \uparrow 0.5} J\left(\widehat{\mathbf{w}}, L, \ell, \frac{(t+\kappa) - E\Pi^q_{agg}(\cdot)}{\left(\operatorname{Var}\Pi^q_{agg}(\cdot)\right)^{1/2}}\right) - J\left(\widehat{\mathbf{w}}, L, \ell, \frac{(t-0.5) - E\Pi^q_{agg}(\cdot)}{\left(\operatorname{Var}\Pi^q_{agg}(\cdot)\right)^{1/2}}\right),$$

with $\widehat{w}_j = \frac{[\mathbf{w}_{agg}]_j - EW_{agg}}{\sqrt{L \operatorname{Var} W_{agg}}}$.

We consider three different examples that characterize possible changes in net assets for individual financial institutions and the overall financial system given a particular stress test class:

Example 3.9 *Our class of stress tests is formed from the following category of risk: credit risk. Specifically, forty percent of all unique AAA-rated mortgage-backed securities have been downgraded to a CCC rating. As a result, the price of each affected mortgage-backed security has declined 80 percent from its original level. In addition, the euro has depreciated by 15 percent relative to the U.S. dollar. We are interested in the possible changes in net assets for each individual financial institution i,* $\forall i \in \{1, \ldots, M\}$*, and we are interested in the possible changes in net assets for the overall financial system.*

Here we have $Q = 1$. Let scalar γ_i be the change in the market value of net assets for financial institution i, $\forall i \in \{1, \ldots, M\}$, when the euro depreciates 15 percent relative to the U.S. dollar. Let scalar γ_{agg} be the change in the market value of net assets for the entire financial system when the euro depreciates 15 percent relative to the U.S. dollar. Random variable $\Pi^1_i(\cdot)$ captures possible changes in net assets for financial institution i arising from the downgrade of AAA-rated mortgage-backed securities, and random variable $\Pi^1_{agg}(\cdot)$ captures possible changes in net assets for the overall financial system arising from this downgrade of mortgage-backed securities. We then have:

$$\Pi_i^{class} = \Pi_i^1(\cdot) + \gamma_i, \forall i \in \{1, \ldots, M\}, \text{ and}$$
$$\Pi_{agg}^{class} = \Pi_{agg}^1(\cdot) + \gamma_{agg}.$$

Random variables $\Pi_i^1(\cdot)$ and $\Pi_{agg}^1(\cdot)$ are constructed by following Example 3.6. We can construct $G_{\Pi_i^1(\cdot)}(t)$ and $G_{\Pi_{agg}^1(\cdot)}(t)$ via asymptotic expansion. With

$$G_{\Pi_i^{class}}(t) = G_{\Pi_i^1(\cdot)}(t - \gamma_i), \forall i \in \{1, \ldots, M\}, \text{ and}$$
$$G_{\Pi_{agg}^{class}}(t) = G_{\Pi_{agg}^1(\cdot)}(t - \gamma_{agg}),$$

we can strongly approximate $G_{\Pi_i^{class}}(t)$ and $G_{\Pi_{agg}^{class}}(t)$ by asymptotic expansion as well. With L total initial AAA-rated mortgage-backed securities, the number of stress test scenarios in this class is $|E^{class}| = \binom{L}{0.4L}$. Corresponding to all of these stress tests are probability distributions that capture possible changes in net assets for each individual financial institution and possible changes in net assets for the overall financial system.

Example 3.10 *Our class of stress tests is formed from two categories of risk:*
(1) Credit risk. Specifically, forty percent of all unique AAA-rated mortgage-backed securities have been downgraded to a CCC rating. As a result, the price of each affected mortgage-backed security has declined 80 percent from its original level.
(2) Exchange rate risk. Specifically, one foreign currency has depreciated by 15 percent relative to the U.S. dollar.

We are interested in the possible changes in net assets for each individual financial institution i, $\forall i \in \{1, \ldots, M\}$, and we are interested in the possible changes in net assets for the overall financial system.

Here we have $Q = 2$. Random variable $\Pi_i^1(\cdot)$ captures possible changes in net assets for financial institution i arising from the downgrade of AAA-rated mortgage-backed securities, and random variable $\Pi_{agg}^{1}(\cdot)$ captures possible changes in net assets for the overall financial system arising from this downgrade of mortgage-backed securities. We can construct $\Pi_{i}^{1}(\cdot)$ and $\Pi_{agg}^{1}(\cdot)$ by following Example 3.6. We can also strongly approximate $G_{\Pi_{i}^{1}(\cdot)}(t)$ and $G_{\Pi_{agg}^{1}(\cdot)}(t)$ via asymptotic expansion. Now, random variable $\Pi_{i}^{2}(\cdot)$ captures possible changes in net assets for financial institution i arising from foreign currency depreciation, and random variable $\Pi_{agg}^{2}(\cdot)$ captures possible changes in net assets for the overall financial system arising from foreign currency depreciation. We can construct $\Pi_{i}^{2}(\cdot)$ and $\Pi_{agg}^{2}(\cdot)$ by following Example 3.5, and as discussed in that example, we can solve for $G_{\Pi_{i}^{2}(\cdot)}(t)$ and $G_{\Pi_{agg}^{2}(\cdot)}(t)$, and their corresponding PMFs, exactly. We then have:

$$\Pi_{i}^{class} = \Pi_{i}^{1}(\cdot) + \Pi_{i}^{2}(\cdot), \forall i \in \{1, \ldots, M\}, \text{ and}$$
$$\Pi_{agg}^{class} = \Pi_{agg}^{1}(\cdot) + \Pi_{agg}^{2}(\cdot), \text{ with}$$

$$g_{\Pi_{i}^{class}}(t) = \left(g_{\Pi_{i}^{1}(\cdot)} * g_{\Pi_{i}^{2}(\cdot)}\right)(t), \forall i \in \{1, \ldots, M\}, \text{ and}$$
$$g_{\Pi_{agg}^{class}}(t) = \left(g_{\Pi_{agg}^{1}(\cdot)} * g_{\Pi_{agg}^{2}(\cdot)}\right)(t).$$

We can generate $g_{\Pi_{i}^{1}(\cdot)}(t)$ and $g_{\Pi_{agg}^{1}(\cdot)}(t)$ from the asymptotic expansions that strongly approximate the CDFs $G_{\Pi_{i}^{1}(\cdot)}(t)$ and $G_{\Pi_{agg}^{1}(\cdot)}(t)$. With L_1 total unique initial AAA-rated mortgage-backed securities and L_2 total foreign currencies, the number of distinct stress test scenarios in this class is $\left|E^{class}\right| = \binom{L_1}{0.4L_1} \times L_2$.

Example 3.11 *Our class of stress tests is formed from three categories of risk:*
(1) Credit risk. Specifically, forty percent of all unique AAA-rated mortgage-backed securities have been downgraded to a CCC rating. As a result, the price of each affected mortgage-backed security has declined 80 percent from its original level.

(2) Exchange rate risk. Specifically, one foreign currency has depreciated by 15 percent relative to the U.S. dollar.

(3) Sovereign risk. Specifically, one country has a writedown of its sovereign debt, which causes the price of each sovereign bond for that country to decrease 10 U.S. dollars.

We are interested in the possible changes in net assets for each individual financial institution i, $\forall i \in \{1, \ldots, M\}$, and we are interested in the possible changes in net assets for the overall financial system.

Here we have $Q = 3$. We build on Example 3.10. Random variable $\Pi_i^3(\cdot)$ captures possible changes in net assets for financial institution i arising from a writedown of sovereign debt, and $\Pi_{agg}^3(\cdot)$ captures possible changes in net assets for the overall financial system arising from a writedown of sovereign debt. We can construct $\Pi_i^3(\cdot)$ and $\Pi_{agg}^3(\cdot)$ by following Example 3.4, and as discussed in that example, we can solve for $G_{\Pi_i^3(\cdot)}(t)$ and $G_{\Pi_{agg}^3(\cdot)}(t)$, and their corresponding PMFs, exactly. We then have

$$\Pi_i^{class} = \Pi_i^1(\cdot) + \Pi_i^2(\cdot) + \Pi_i^3(\cdot), \forall i \in \{1, \ldots, M\}, \text{ and}$$
$$\Pi_{agg}^{class} = \Pi_{agg}^1(\cdot) + \Pi_{agg}^2(\cdot) + \Pi_{agg}^3(\cdot), \text{ with}$$

$$g_{\Pi_i^{class}}(t) = \left(g_{\Pi_i^1(\cdot)} * g_{\Pi_i^2(\cdot)} * g_{\Pi_i^3(\cdot)}\right)(t), \forall i \in \{1, \ldots, M\}, \text{ and}$$
$$g_{\Pi_{agg}^{class}}(t) = \left(g_{\Pi_{agg}^1(\cdot)} * g_{\Pi_{agg}^2(\cdot)} * g_{\Pi_{agg}^3(\cdot)}\right)(t).$$

With L_1 total unique initial AAA-rated mortgage-backed securities, L_2 total foreign currencies, and L_3 total countries that have issued sovereign bonds, the total number of stress test scenarios in this class is $\left|E^{class}\right| = \binom{L_1}{0.4L_1} \times L_2 \times L_3$. Corresponding to this class of stress tests with its many underlying stress test scenarios is a probability distribution summarizing possible balance sheet effects for each individual financial institution and a probability distribution summarizing possible balance sheet effects for the overall financial system.

3.3 Conclusion

The global financial crisis of 2008 forced a regulatory paradigm shift for central banks and other supervisory institutions around the world. As a result of the global financial crisis, supervisory institutions expanded their mandates. They rethought their specific functions, and they developed new approaches to regulation. In particular, central banks such as the Federal Reserve began to develop and implement stress tests as part of a new financial stability mandate, and they nominally shifted their regulatory approaches from ones that were purely microprudential to ones that were both macroprudential and microprudential.

Within the United States, even though stress tests were used in the financial industry prior to the global financial crisis, annual supervisory stress tests only became a mandatory part of the Federal Reserve's regulatory toolkit with the passage of the 2010 Dodd-Frank Act. Globally, it was the post-crisis Basel III capital framework that articulated principles of stress testing for regulatory institutions; these global institutions could then use the Basel III framework as a constructive set of guidelines for developing their own stress tests.

The Federal Reserve's set of stress tests, while designed to assess the stability of individual financial institutions and the financial system as a whole, can be made much more comprehensive and impactful. We want the Federal Reserve's stress testing tools to be maximally informative, but in their current form, they only provide a limited view of the financial landscape and the broader macroeconomy. The number of stress tests that the Federal Reserve executes annually is very small, and the stress test scenarios are often calibrated to past historical events. As a result, current stress tests do not adequately assess the financial system's overall health and ability to maintain operations and obligations in the presence of a broad range of possible negative shocks. The financial system can potentially be ill-equipped to handle certain realistic stressed scenarios, but this weakness will never be uncovered while employing the Federal Reserve's current stress testing approach. The present work shows how to massively increase the total number of stress test scenarios without increasing the computational burden; this work therefore substantially strengthens the Federal Reserve's stress testing process. Rather than collecting a small number of data points summarizing financial institutions' balance sheet effects for each individual stress test, the Federal Reserve can instead construct entire probability distributions that capture financial institutions' balance sheet effects for a *class* of stress tests. Associated with each class of stress tests is a corresponding probability distribution for each individual financial institution and a probability distribution for the financial system as a whole.

The approach to stress tests in this work differs from the Federal Reserve's existing approach. The Federal Reserve currently develops a very small number of stress tests that generally mimic past historical events. The present work, meanwhile, articulates a different framework for the Federal Reserve; according to the present work, the Federal Reserve would first identify different classes of stress tests, and then within each class, the Federal Reserve would generate an exhaustive list of constituent stress tests. The Federal Reserve would form each class of stress tests by identifying certain categories of risk. Different stress tests within the same class have the same categories of risk, but the manner by which such risks manifest themselves within the financial system would differ. The approach of this work for generating large classes of stress tests is top-down and macroprudential in spirit;

we are assuming that the financial system inherently has certain types of overall stressors, but we are agnostic to how these stressors ultimately appear in the system. We consider *all* of the multitudinous ways by which these stressors can potentially manifest themselves.

Going forward, it is important that we craft a methodology for how to construct classes of stress tests. We would want to know which categories of risk are the most relevant and therefore appropriate for more thorough examination. We would also want to know how many different classes of stress tests are sufficient to enable thorough analysis of the financial system. Having the correct set of tools is crucial for the Federal Reserve and other central banks. These tools, when properly designed and employed, provide regulatory institutions with the ability to pinpoint potential sources of weakness within the financial system, evaluate the financial system's overall health, and more generally survey the financial landscape. The financial system is inherently global, and shocks within the financial system transmit to the broader macroeconomy, so it is extremely important that we develop and employ the right set of regulatory tools.

References

ACEMOGLU, D., AKCIGIT, U. and KERR, W. (2016). Networks and the macroeconomy: An empirical exploration. *NBER Macroeconomics Annual*, **30** (1), 273–335.

—, CARVALHO, V. M., OZDAGLAR, A. and TAHBAZ-SALEHI, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, **80** (5), 1977–2016.

—, DAHLEH, M. A., LOBEL, I. and OZDAGLAR, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, **78** (4), 1201–1236.

—, OZDAGLAR, A. and PARANDEHGHEIBI, A. (2010). Spread of (mis) information in social networks. *Games and Economic Behavior*, **70** (2), 194–227.

—, — and TAHBAZ-SALEHI, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, **105** (2), 564–608.

—, — and — (2017). Microeconomic origins of macroeconomic tail risks. *American Economic Review*, **107** (1), 54–108.

ACHARYA, S., BENHABIB, J. and HUO, Z. (2017). The anatomy of sentiment-driven fluctuations. *Working Paper*.

ACHARYA, V. (2009). A theory of systemic risk and design of prudential bank regulation. *Journal of Financial Stability*, **5** (3), 224–255.

—, ENGLE, R. and PIERRET, D. (2014). Testing macroprudential stress tests: The risk of regulatory risk weights. *Journal of Monetary Economics*, **65**, 36–53.

AFONSO, G., KOVNER, A. and SCHOAR, A. (2011). Stressed, not frozen: The federal funds market in the financial crisis. *The Journal of Finance*, **66** (4), 1109–1139.

AHLSWEDE, R. and KATONA, G. O. (1978). Graphs with maximal number of adjacent pairs of edges. *Acta Mathematica Hungarica*, **32** (1-2), 97–120.

ALLEN, F. and BABUS, A. (2009). Networks in finance. *The network challenge: Strategy, profit, and risk in an interlinked world*.

ANDERSON, R. W. (ed.) (2016). *Stress testing and macroprudential regulation: A transatlantic assessment*. CEPR Press.

Angeletos, G.-M., Collard, F. and Dellas, H. (2017). An anatomy of the business cycle. *Working Paper*.

— and La'O, J. (2013). Sentiments. *Econometrica*, **81** (2), 739–779.

Bak, P., Chen, K., Scheinkman, J. and Woodford, M. (1993). Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche*, **47** (1), 3–30.

Ballester, C., Calvó-Armengol, A. and Zenou, Y. (2006). Who's who in networks. Wanted: The key player. *Econometrica*, **74** (5), 1403–1417.

Banerjee, A., Breza, E., Chandrasekhar, A. G. and Mobius, M. (2016). Naive learning with uninformed agents. *Working Paper*.

Baqaee, D. (2013). Labor intensity in an interconnected economy. *Working Paper*.

Baqaee, D. R. and Farhi, E. (2018). The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem. *Working Paper*.

Barrot, J.-N. and Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, **131** (3), 1543–1592.

Barsky, R. B. and Sims, E. R. (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review*, **102** (4), 1343–77.

Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B. and Stiglitz, J. E. (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control*, **36** (8), 1121–1141.

Benhabib, J., Wang, P. and Wen, Y. (2015). Sentiments and aggregate demand fluctuations. *Econometrica*, **83** (2), 549–585.

Bisgaard, M., Sønderskov, K. M. and Dinesen, P. T. (2016). Reconsidering the neighborhood effect: Does exposure to residential unemployment influence voters' perceptions of the national economy? *The Journal of Politics*, **78** (3), 719–732.

Bisin, A., Horst, U. and Özgür, O. (2006). Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory*, **127** (1), 74–116.

Björklund, A. and Salvanes, K. G. (2011). Education and family background: Mechanisms and policies. In *Handbook of the Economics of Education*, vol. 3, Elsevier, pp. 201–247.

Boehm, C., Flaaen, A. and Pandalai-Nayar, N. (Forthcoming). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake. *The Review of Economics and Statistics*.

Bookstaber, R., Cetina, J., Feldberg, G., Flood, M. and Glasserman, P. (2014). Stress tests to promote financial stability: Assessing progress and looking to the future. *Journal of Risk Management in Financial Institutions*, **7** (1), 16–25.

Borchgrevink, H., Ellingsrud, S. and Hansen, F. (2014). Macroprudential regulation–What, why and how? *Norges Bank Staff Memo*.

Borio, C. (2003). Towards a macroprudential framework for financial supervision and regulation? *CESifo Economic Studies*, **49** (2), 181–215.

—, Drehmann, M. and Tsatsaronis, K. (2014). Stress-testing macro stress testing: Does it live up to expectations? *Journal of Financial Stability*, **12**, 3–15.

Brauer, A. (1952). Limits for the characteristic roots of a matrix. IV: Applications to stochastic matrices. *Duke Mathematical Journal*, **19** (1), 75–91.

Brock, W. A. and Durlauf, S. N. (2001a). Discrete choice with social interactions. *The Review of Economic Studies*, **68** (2), 235–260.

— and — (2001b). Interactions-based models. In *Handbook of Econometrics*, vol. 5, Elsevier, pp. 3297–3380.

Caccioli, F., Shrestha, M., Moore, C. and Farmer, J. D. (2014). Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance*, **46**, 233–245.

Calvó-Armengol, A., Patacchini, E. and Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, **76** (4), 1239–1267.

— and Zenou, Y. (2004). Social networks and crime decisions: The role of social structure in facilitating delinquent behavior. *International Economic Review*, **45** (3), 939–958.

Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, **28** (4), 23–48.

Chandrasekhar, A. G., Larreguy, H. and Xandri, J. P. (2018). Testing models of social learning on networks: Evidence from two experiments. *Working Paper*.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, **104** (11), 3600–3634.

Chodorow-Reich, G. (2018). Geographic cross-sectional fiscal spending multipliers: What have we learned? *Working Paper*.

Christiano, L., Eichenbaum, M. and Rebelo, S. (2011). When is the government spending multiplier large? *Journal of Political Economy*, **119** (1), 78–121.

Claessens, S. (2015). An overview of macroprudential policy tools. *Annual Review of Financial Economics*, **7**, 397–422.

Clark, A. and Large, A. (2011). *Macroprudential policy: Addressing the things we don't know*. Group of Thirty.

Clement, P. (2010). The term 'macroprudential': Origins and evolution. *BIS Quarterly Review*, p. 59.

Conlisk, J. (1985). Comparative statics for Markov chains. *Journal of Economic Dynamics and Control*, **9** (2), 139–151.

Cont, R., Moussa, A. and Santos, E. (2013). Network structure and systemic risk in banking systems. *Handbook on Systemic Risk*, pp. 327–336.

Cooper, C. and Frieze, A. (2012). Stationary distribution and cover time of random walks on random digraphs. *Journal of Combinatorial Theory, Series B*, **102**, 329–362.

d'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *The American Economic Review*, **78** (5), 1133–1137.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, **69** (345), 118–121.

DeMarzo, P. M., Vayanos, D. and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, **118** (3), 909–968.

Demekas, D. (2015). Designing effective macroprudential stress tests: Progress so far and the way forward. *IMF Working Paper*.

Eger, S. (2016a). Opinion dynamics and wisdom under out-group discrimination. *Mathematical Social Sciences*, **80**, 97–107.

— (2016b). On limits of powers of certain absolutely row-stochastic matrices. *Linear Algebra and its Applications*, **508**, 1–13.

Eggertsson, G. B. (2011). What fiscal policy is effective at zero interest rates? *NBER Macroeconomics Annual*, **25** (1), 59–112.

Elliott, M., Golub, B. and Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, **104** (10), 3115–3153.

Epple, D. and Romano, R. E. (2011). Peer effects in education: A survey of the theory and evidence. In *Handbook of Social Economics*, vol. 1, Elsevier, pp. 1053–1163.

Erdös, P. and Rényi, A. (1959). On the central limit theorem for samples from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci*, **4**, 49–61.

Farboodi, M. (2014). Intermediation and voluntary exposure to counterparty risk. *Working Paper*.

Farhi, E. and Werning, I. (2016). Fiscal multipliers: Liquidity traps and currency unions. In *Handbook of Macroeconomics*, vol. 2, Elsevier, pp. 2417–2492.

Farmer, R. and Guo, J.-T. (1994). Real business cycles and the animal spirits hypothesis. *Journal of Economic Theory*, **63** (1), 42–72.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, **79** (3), 733–772.

Gai, P. and Kapadia, S. (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **466** (2120), 2401–2423.

Gale, D. and Kariv, S. (2003). Bayesian learning in social networks. *Games and Economic Behavior*, **45** (2), 329–346.

Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F. and Yariv, L. (2010). Network games. *The Review of Economic Studies*, **77** (1), 218–244.

Glaeser, E. L., Sacerdote, B. I. and Scheinkman, J. A. (2003). The social multiplier. *Journal of the European Economic Association*, **1** (2-3), 345–353.

Glasserman, P. and Tangirala, G. (2016). Are the Federal Reserve's stress test results predictable? *The Journal of Alternative Investments*, **18** (4), 82–97.

Goidel, R. K. and Langley, R. E. (1995). Media coverage of the economy and aggregate economic evaluations: Uncovering evidence of indirect media effects. *Political Research Quarterly*, **48** (2), 313–328.

Golub, B. and Jackson, M. O. (2007). Naive learning in social networks: Convergence, influence and wisdom of crowds. *Working Paper*.

— and — (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, **2** (1), 112–49.

Goyal, S. and Moraga-Gonzalez, J. L. (2001). R&D networks. *Rand Journal of Economics*, pp. 686–707.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, **83** (6), 1420–1443.

Greenlaw, D., Kashyap, A., Schoenholtz, K. and Shin, H. S. (2012). Stressed out: Macroprudential principles for stress testing. *Chicago Booth Research Paper*, (12-08).

GRUNDKE, P. (2011). Reverse stress tests with bottom-up approaches. *The Journal of Risk Model Validation*, **5** (1), 71.

GUALDI, S., CIMINI, G., PRIMICERIO, K., DI CLEMENTE, R. and CHALLET, D. (2016). Statistically validated network of portfolio overlaps and systemic risk. *Scientific Reports*, **6**, 1–14.

HAGEDORN, M., MANOVSKII, I. and MITMAN, K. (2018). The fiscal multiplier. *Working Paper*.

HANSON, S. G., KASHYAP, A. K. and STEIN, J. C. (2011). A macroprudential approach to financial regulation. *Journal of Economic Perspectives*, **25** (1), 3–28.

HAREL, M., MOSSEL, E., STRACK, P. and TAMUZ, O. (2017). Groupthink and the failure of information aggregation in large groups. *Working Paper*.

HEALY, A. and LENZ, G. S. (2017). Presidential voting and the local economy: Evidence from two population-based data sets. *The Journal of Politics*, **79** (4), 1419–1432.

HECKMAN, J. J. and SNYDER, J. M. (1997). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *The RAND Journal of Economics*, **28**, S142–S189.

HETHERINGTON, M. J. (1996). The media's role in forming voters' national economic evaluations in 1992. *American Journal of Political Science*, pp. 372–395.

HIRTLE, B. and LEHNERT, A. (2015). Supervisory stress tests. *Annual Review of Financial Economics*, **7**, 339–355.

HÖGLUND, T. (1978). Sampling from a finite population. A remainder term estimate. *Scandinavian Journal of Statistics*, **5** (1), 69–71.

HORST, U. and SCHEINKMAN, J. A. (2006). Equilibria in systems of social interactions. *Journal of Economic Theory*, **130** (1), 44–77.

HORVATH, M. (1998). Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics*, **1** (4), 781–808.

— (2000). Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, **45** (1), 69–106.

HUO, Z. and TAKAYAMA, N. (2015). Higher order beliefs, confidence, and business cycles. *Working Paper*.

ILZETZKI, E., MENDOZA, E. G. and VÉGH, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, **60** (2), 239–254.

Jackson, M. O., Rogers, B. W. and Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, **55** (1), 49–95.

— and Yariv, L. (2011). Diffusion, strategic interaction, and social structure. In *Handbook of Social Economics*, vol. 1, Elsevier, pp. 645–678.

— and Zenou, Y. (2015). Games on networks. In *Handbook of Game Theory with Economic Applications*, vol. 4, Elsevier, pp. 95–163.

Kashyap, A. K., Berner, R. and Goodhart, C. A. (2011). The macroprudential toolkit. *IMF Economic Review*, **59** (2), 145–161.

Kemeny, J. G. and Snell, J. L. (1960). *Finite Markov chains*. Princeton, NJ: D. Van Nostrand Company.

König, M., Liu, X. and Zenou, Y. (Forthcoming). R&D networks: Theory, empirics and policy implications. *Review of Economics and Statistics*.

Levy-Carciente, S., Kenett, D. Y., Avakian, A., Stanley, H. E. and Havlin, S. (2015). Dynamical macroprudential stress testing using network theory. *Journal of Banking & Finance*, **59**, 164–181.

Liebeg, D. and Posch, M. (2011). Macroprudential regulation and supervision: From the identification of systemic risks to policy measures. *Financial Stability Report*, **21**, 67–83.

Markose, S., Giansante, S. and Shaghaghi, A. R. (2012). 'Too interconnected to fail' financial network of US CDS market: Topological fragility and systemic risk. *Journal of Economic Behavior & Organization*, **83** (3), 627–646.

Marotta, L., Miccichè, S., Fujiwara, Y., Iyetomi, H., Aoyama, H., Gallegati, M. and Mantegna, R. N. (2015). Bank-firm credit network in Japan: An analysis of a bipartite network. *PLOS One*, **10** (5), 1–18.

McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*, *Monographs on Statistics and Applied Probability*, vol. 37. Chapman and Hall, 2nd edn.

Milani, F. (2017). Sentiment and the US business cycle. *Journal of Economic Dynamics and Control*, **82**, 289–311.

Nakamura, E. and Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review*, **104** (3), 753–92.

Nolan, J. P. (2014). Financial modeling with heavy-tailed stable distributions. *Wiley Interdisciplinary Reviews: Computational Statistics*, **6** (1), 45–55.

Oberfield, E. (2018). A theory of input–output architecture. *Econometrica*, **86** (2), 559–589.

Petrella, G. and Resti, A. (2013). Supervisors as information producers: Do stress tests reduce bank opaqueness? *Journal of Banking & Finance*, **37** (12), 5406–5420.

Robinson, J. (1978). An asymptotic expansion for samples from a finite population. *The Annals of Statistics*, **6** (5), 1005–1011.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, vol. 3, Elsevier, pp. 249–277.

Scheinkman, J. A. and Woodford, M. (1994). Self-organized criticality and economic fluctuations. *The American Economic Review*, **84** (2), 417–421.

Schuermann, T. (2014). Stress testing banks. *International Journal of Forecasting*, **30** (3), 717–728.

Seneta, E. (1981). *Non-negative matrices and Markov chains*. Springer Series in Statistics, Springer, New York.

Suzumura, K. (1992). Cooperative and noncooperative R&D in an oligopoly with spillovers. *The American Economic Review*, pp. 1307–1320.

Topa, G. (2001). Social interactions, local spillovers and unemployment. *The Review of Economic Studies*, **68** (2), 261–295.

Wall, L. (2014). Measuring capital adequacy: Supervisory stress tests in a Basel world. *Journal of Financial Perspectives*, **2** (1), 85–94.

Williams, J. C. (2015). Macroprudential policy in a microprudential world. *FRBSF Economic Letter*, (18), 1–7.

Woodford, M. (2011). Simple analytics of the government expenditure multiplier. *American Economic Journal: Macroeconomics*, **3** (1), 1–35.

Zawadowski, A. (2013). Entangled financial systems. *The Review of Financial Studies*, **26** (5), 1291–1323.

Appendix A

Appendix to Chapter 1

A.1 Data

A.1.1 Construction of Viewership, Listenership, and Readership Statistics

IndieWire provides nightly primetime television viewership statistics in 2016 for almost every ad-supported broadcast and cable network.[1] Only networks that provide non-sports news and/or talk coverage are included in the sample. PBS network viewership is set equal to 2016 total average viewership for its NewsHour program.[2]

Radio listenership data comes from Wikipedia[3], with many of the statistics

[1] http://www.indiewire.com/2016/12/cnn-fox-news-msnbc-nbc-ratings-2016-winners-losers-1201762864/. Accessed February 21, 2018.

[2] Statistic provided by the Pew Research Center. http://www.journalism.org/chart/pbs-newshour-viewership/. Accessed February 21, 2018.

[3] "List of Most-Listened-To Radio Programs." https://en.wikipedia.org/wiki/List_of_most-listened-to_radio_programs. Accessed February 21, 2018.

originating in TALKERS magazine.[4] The data consists of weekly listenership for several radio programs, mostly from 2017. I assume that daily listenership equals weekly listenership, and I only include those radio shows categorized as news/talk radio.

Readership statistics for magazines, newspapers, business journals, and business publications come from the Alliance for Audited Media, Media Intelligence Center.[5] To obtain readership statistics for magazines, I apply the following filters: Status - Active, Publication - Magazine, and Country - United States. I then restrict the set of magazines to ones with the following SRDS Classifications for Magazines: "News", "Political & Social Topics", "Business", "Newspaper: Alternative", "Popular Culture", "General Editorial/Content", and "Metropolitan/Regional/State". To obtain readership statistics for newspapers, the filters that I apply are as follows: Status - Active, Publication - Newspaper, and Country - United States. To obtain readership statistics for business journals and publications, the filters that I apply are as follows: Status - Active, Publication - Business, and Country - United States. Readership statistics were mostly collected in 2016 and 2017. I remove Costco Connection from the magazine sample. For those cases in which a publication has separate statistics for its weekday circulation and its Saturday/Sunday circulation, I designate the larger value as that publication's readership.

CBS has the largest television network nightly primetime audience, with approximately 8.8 million viewers; NPR's *Morning Edition*, NPR's *All Things Considered*, and APM's *Marketplace* have the largest weekly radio listening audience, each with approximately 14.6 million listeners; *USA Today* has the largest newspaper subscriber base, with a weekday circulation of approximately 2.1 million; *People* magazine has

[4]http://www.talkers.com/top-talk-audiences/. Archived version from September 6, 2017.

[5]Accessed February 21, 2018.

the largest magazine subscriber base, with a circulation of approximately 3.4 million; and *The Who's Who in Building & Construction* has the largest business journal and business publication readership, with a circulation of approximately 512,000.

There is a total of 1867 news/talk media sources. The audience for each of these sources is culled from the entire U.S. population. Since the voting population is just a subset of the entire U.S. population, we must scale down the viewership, listenership, and readership numbers to reflect the relatively smaller size of the voting population. We assume that the entire pool of potential consumers of such news/talk media consists of 250,293,421 individuals, the projected 2016 population size for people aged 18 and over.[6] We then multiply each audience size by the factor 137.5×10^6 / 250,293,421 to obtain audience statistics for the voting population.

A.1.2 Summary Statistics

We present summary statistics for the base graph, the media-originating graph, and the composite graph that pools both the base and media-originating linkages. In the base graph, with its undirected edges, the average degree is 51.0 with a standard deviation of 7.07. The minimum degree is 17 and the maximum degree is 96. There is a total of 3,575,017,297 undirected edges. In the graph with directed media-originating linkages, the average out-degree is 19.7 with a standard deviation of 17.0. The minimum out-degree is zero, and the maximum out-degree is 165. The average in-degree is 19.7 with a standard deviation of 8,633.3. The minimum in-degree is zero, and the maximum in-degree is 8,020,586, the largest audience size among all media sources. The total number of directed edges in this graph

[6]Table 1. Projected Population by Single Year of Age, Sex, Race, and Hispanic Origin for the United States: 2014 to 2060, from "2014 National Population Projections Dataset," United States Census Bureau. https://www.census.gov/data/datasets/2014/demo/popproj/2014-popproj.html.

is 2,712,493,694. Agents' out-degrees occur in multiples of 15 because, if they are exposed to one media source, they become connected to 15 individuals through the reporting that the media source provides. The average and median number of media sources that people are exposed to are, respectively, 1.32 and 1. 28,003 distinct people are featured in employment-related news stories over the course of 15 weeks. In the composite graph, that is, the graph that pools edges from the base graph and the media-originating graph, the average out-degree is 70.7 with a standard deviation of 18.4, the minimum out-degree is 17, and the maximum out-degree is 230. The average in-degree is 70.7 with a standard deviation of 8,633.3, the minimum in-degree is 17, and the maximum in-degree is 8,020,651. We observe that the counter-cumulative distribution function of in-degrees for the total network (Figure 1.7, bottom) directly incorporates the distributional features of the counter-cumulative distribution function of degrees for the base graph (Figure 1.7, top left) and the counter-cumulative distribution function of in-degrees for the media graph (Figure 1.7, middle).

A.2 Section 1.3 Supplemental Material

First, we assume that agents' observation network is just the base graph. We would like to construct the distribution of possible average local unemployment rates given that there is an overall 9.6-percent global unemployment rate. We begin by computing the vector of average weighted in-degrees, where we assume that each agent assigns an equal weight to his or her observations of employment status. In this setting, the vector of average weighted in-degrees is the relevant network-derived vector of agent weights. With agent weights summing to 1, the average weighted in-degree is 7.27×10^{-9} with a standard deviation of 9.98×10^{-10}, the minimum average weighted in-degree is 2.67×10^{-9}, the maximum average weighted in-degree is 1.37×10^{-8}, and the median average weighted in-degree is 7.25×10^{-9}. From these average weighted in-degrees, we can compute each agent's effective representation in the population. On average, each agent effectively represents 1 agent. The agent with the smallest weight effectively represents 0.367 agents. The agent with the largest weight effectively represents 1.89 agents, and the agent with the median weight effectively represents 0.996 agents. There is not much heterogeneity in individual agents' effective representations.

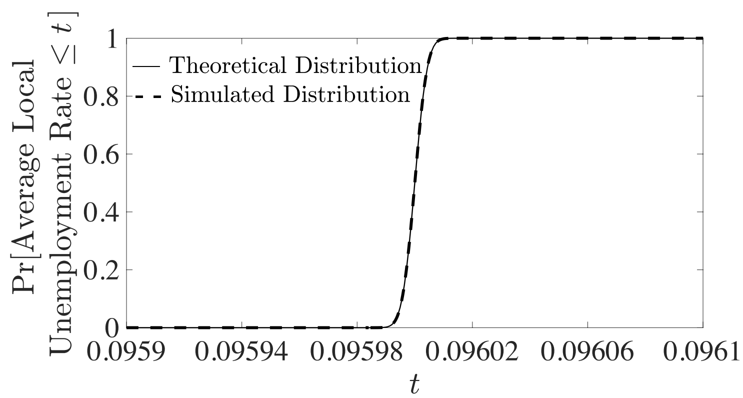


Figure A.1: *Distribution of the average local unemployment rate, $f = 0.096$, assuming that configurations of unemployment in the economy are equally likely, agents' social observation network is the base graph, and agents assign an equal weight to each of their out-linkages.*

Figure A.1 plots the distribution of possible average local unemployment rates when the global unemployment rate is 0.096, assuming that each configuration of unemployment in the economy is equally likely to occur and agents' observation network is the base graph. The theoretical CDF, constructed using Theorem 1.13 in Section 1.6, overlays an empirical CDF. The empirical CDF is constructed by simulation, randomly drawing 100,000 configurations from the set of all possible configurations and computing the associated average local unemployment rate for each configuration. The theoretical and empirical mean of this distribution is 0.096 and the theoretical standard deviation is 3.45×10^{-6}. These quantities are computed from Theorems 1.8 and 1.9 in Section 1.6. The minimum possible average local unemployment rate is 0.0735, and the maximum possible average local unemployment rate is 0.120; both quantities are constructed from Theorem 1.10 in Section 1.6. When agents' observation network is solely the base graph, the average local unemployment rate does not meaningfully vary with configuration. As a result, the election outcome, and consequently the outcome for the economy, does not particularly depend on configuration:

Example A.1 (Voting Outcome, Base Graph) *Aggregate voting behavior is characterized by Equation 1.4. The unemployment rate is 9.6 percent. Given that voters' observation network is the base graph, and voters equally weight each of their observations, the probability that Trump's expected vote share exceeds 0.5 is zero:*

$$\Pr\left[\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right) > 0.10\right] = 0.$$

With certainty, the election outcome favors Clinton.

We next consider a social observation network whose method of construction differs slightly from the network in the main text. The composite social observation network considered here pools linkages from a base graph and a media graph. The media graph is the same as the one in the main text. Meanwhile, the base graph differs; it is constructed by assuming that each agent has a self-loop and, on average, 20 reciprocal linkages rather than 50 reciprocal linkages. Figure A.2 presents the counter-cumulative distribution function of degrees for the base graph. The base graph has an average degree of 21.0 with a standard deviation of 4.47. The minimum degree is 2, while the maximum degree is 50. There is a total of 1,512,512,933 undirected edges. Figure A.2 presents the counter-cumulative distribution functions of in-degrees and out-degrees for the media graph and the composite graph. The summary statistics for the unchanged media graph are

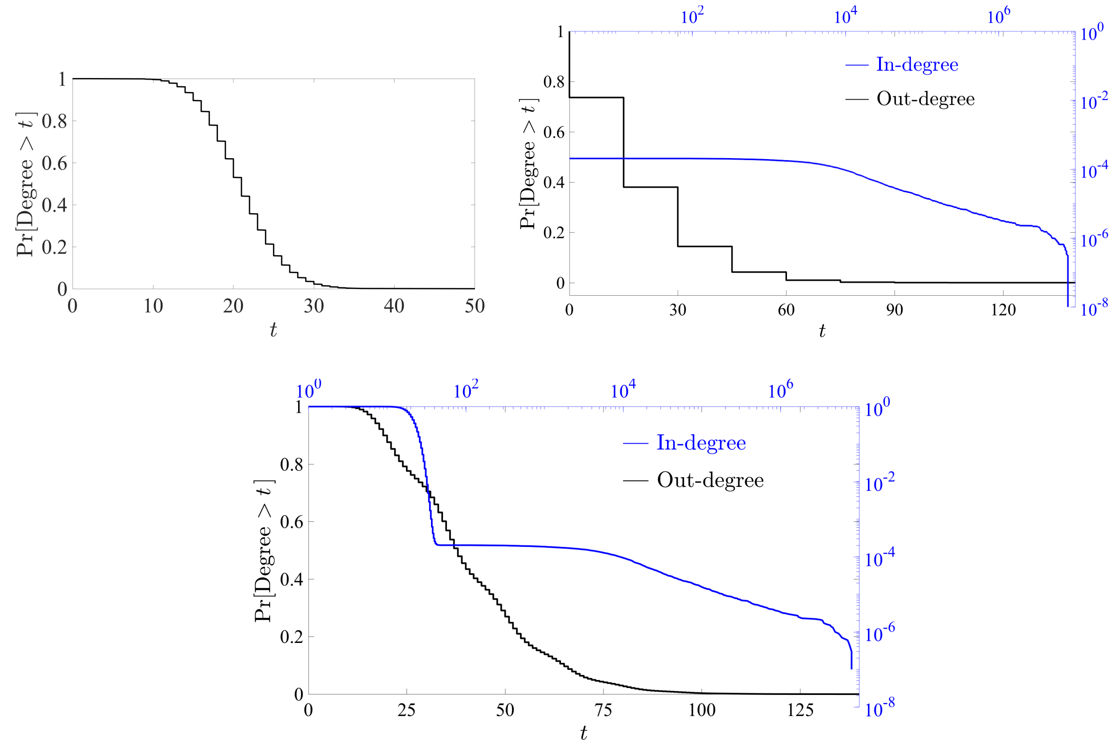


Figure A.2: *Counter-cumulative distribution function (CCDF) of degrees for the base graph when the average number of reciprocal linkages that each agent forms is equal to 20 (top left). CCDFs of out- and in-degrees for the network of media-originating linkages (top right). CCDFs of out- and in-degrees for the resulting composite network (bottom).*

discussed in the main text and Appendix A.1. The composite graph has an average out-degree of 40.7, with a standard deviation of 17.5, a minimum out-degree of 2, and a maximum out-degree of 188. The composite graph also has an average in-degree of 40.7, with a standard deviation of 8633.3, a minimum in-degree of 2, and a maximum in-degree of 8,020,613.

We assume that agents equally weight each of their linkages on the composite graph. We can then compute the vector of agent weights, which is the vector of average weighted in-degrees for the composite graph. On average, each agent has an effective weight of 1 agent. The effective minimum weight is 0.0498 agents, and the effective maximum weight is 160,770.2 agents. The median agent has an effective weight of 0.600 agents. The left side of Figure A.3 plots the counter-cumulative distribution function of average weighted in-degrees. This distribution of agent weights is heavy-tailed.

The distribution of possible average local unemployment rates, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, is on the right side of Figure A.3. The theoretical CDF overlays an empirical CDF, the latter of which is constructed by randomly drawing 100,000 configurations of unemployment from the set of all possible configurations consistent with a 9.6-percent unemployment rate, and then computing the associated average local unemployment rate for each configuration. The theoretical and empirical mean of this distribution is 0.096. The theoretical standard deviation for this distribution is 0.00433, or 0.433 percentage points, and the size of two standard deviations about the distribution's mean value is 1.73 percentage points. Staying within this two-standard-deviation band, the average local unemployment rate can generally vary from 8.73 percent to 10.5 percent. The lowest possible average local unemployment rate is 3.64 percent, and the highest possible average local unemployment rate is 47.6 percent. The probability that the average local unemployment rate exceeds 10 percent is 17.7 percent. Therefore, the probability that the election outcome favors Trump is 17.7 percent, and the probability that the election outcome favors Clinton is 82.3 percent. This particular economy exhibits greater configuration dependence than the economy studied in the main text.

We proceed to consider another social observation network whose method of construction also differs slightly from the network in the main text. As before, the composite social observation network pools linkages from a base graph and a media graph. The base graph is the same as the one in the main text. Meanwhile, the media graph differs; it is constructed by assuming that each news/talk media source publishes five stories on the issue of jobs and unemployment rather than publishing

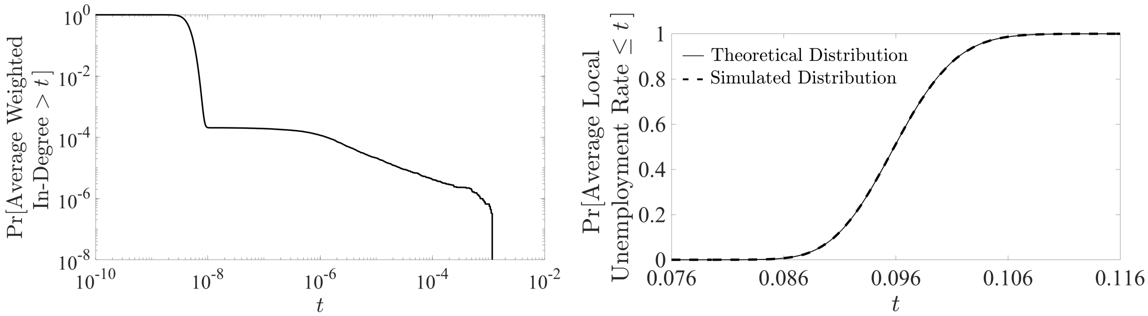


Figure A.3: *Counter-cumulative distribution function of average weighted in-degrees for the composite network, assuming that each agent assigns an equal weight to each of his out-linkages and the base graph consists of agents forming on average 20 reciprocal linkages (left). Distribution of the average local unemployment rate,* $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, *when* $f = 0.096$, *assuming that configurations of unemployment in the economy are equally likely (right).*

15 stories. As a result, for each media source, there are five featured individuals. Figure A.4 presents the counter-cumulative distribution function of degrees for the base graph. The statistical features of the degree distribution for the base graph can be found in the main text and Appendix A.1. Figure A.4 additionally presents the counter-cumulative distribution functions of in-degrees and out-degrees for the media graph and the composite graph. In the media-originating graph, with its 904,164,565 directed edges, the average out-degree is 6.58 with a standard deviation of 5.65, and the average in-degree is 6.58, with a standard deviation of 4,984.5. The counter-cumulative distribution function of out-degrees for the media-originating graph is a step function because agents accumulate 5 out-edges for every media source in which they are an audience member. Therefore, out-degrees for the media-originating graph occur in multiples of 5. Most voters have zero in-degree for the media-originating graph because they are not featured in news/talk media outlets; there are only 9,335 individuals featured in employment-related news stories. In the composite graph, the average out-degree is 57.6 with a standard deviation of 9.05, and the average in-degree is 57.6 with a standard deviation of 4984.5. The

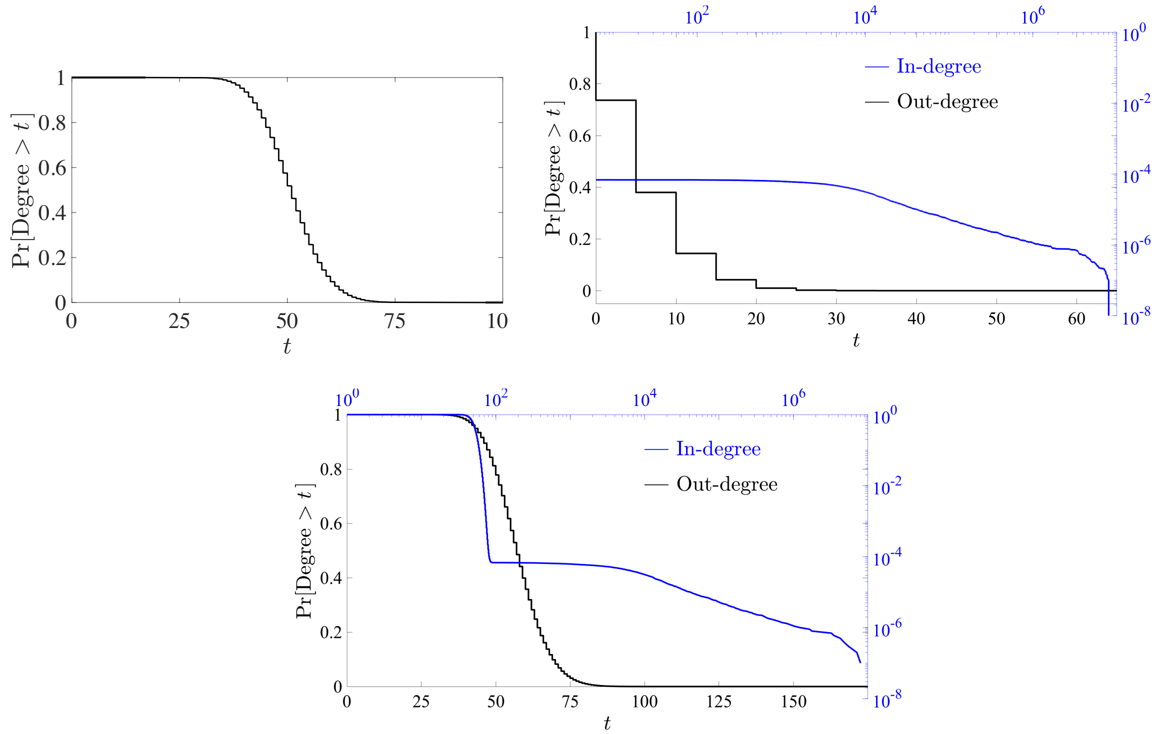


Figure A.4: *Counter-cumulative distribution function (CCDF) of degrees for the base graph (top left). CCDFs of out- and in-degrees for the network of media-originating linkages when each media source publishes five stories on the issue of jobs and unemployment (top right). CCDFs of out- and in-degrees for the composite network (bottom).*

maximum in-degree is 8,020,644.

We assume that agents equally weight each of their linkages on the composite graph. We can then compute the vector of agent weights, which is the vector of average weighted in-degrees for the composite graph. On average, each agent has an effective weight of 1 agent. The effective minimum weight is 0.304 agents, and the effective maximum weight is 131,516.8 agents. The median agent has an effective weight of 0.889 agents. The left side of Figure A.5 plots the counter-cumulative distribution function of average weighted in-degrees. This distribution of agent weights is heavy-tailed.

The distribution of possible average local unemployment rates, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$,

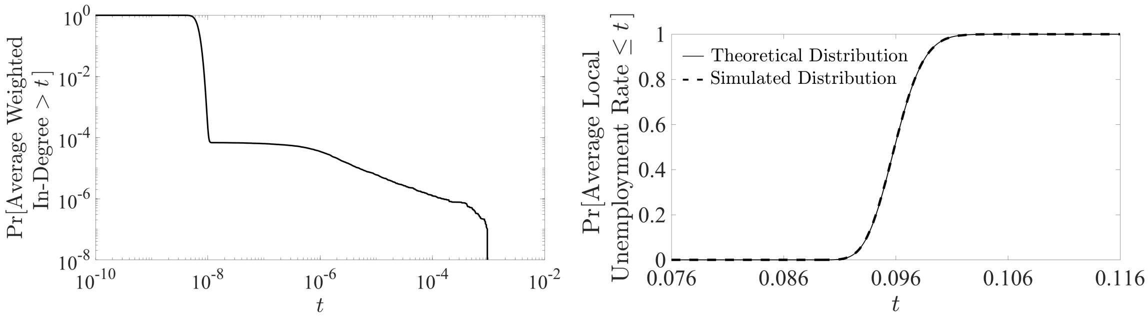


Figure A.5: *Counter-cumulative distribution function of average weighted in-degrees for the composite network, assuming that each agent assigns an equal weight to each of his out-linkages and each media source publishes five stories on the issue of jobs and unemployment (left). Distribution of the average local unemployment rate,* $G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$*, when* $f = 0.096$*, assuming that configurations of unemployment in the economy are equally likely (right).*

is on the right side of Figure A.5. The theoretical CDF overlays an empirical CDF, the latter of which is constructed by randomly drawing 100,000 configurations of unemployment from the set of all possible configurations consistent with a 9.6-percent unemployment rate, and then computing the associated average local unemployment rate for each configuration. The theoretical and empirical mean of this distribution is 0.096. The theoretical standard deviation for this distribution is 0.00205, or 0.205 percentage points, and the size of two standard deviations about the distribution's mean value is 0.820 percentage points. Staying within this two-standard-deviation band, the average local unemployment rate can generally vary from 9.19 percent to 10.0 percent. The lowest possible average local unemployment rate is 6.55 percent, and the highest possible average local unemployment rate is 21.5 percent. The probability that the average local unemployment rate exceeds 10 percent is 3.27 percent. Therefore, the probability that the election outcome favors Trump is 3.27 percent, and the probability that the election outcome favors Clinton is 96.7 percent. This particular economy exhibits less configuration dependence than the economy studied in the main text.

A.3 Section 1.4 Supplemental Theorem

We first characterize $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n) = [\bar{\mathbf{A}}^q]_{i*}\mathbf{b}(N, n)$ for all finite q and in the limit as $q \to \infty$ in terms of the fundamental features of $\bar{\mathbf{A}}$, and we study the rate of convergence of $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$. We then characterize $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}}) = [\bar{\mathbf{A}}^q]_{i*}$ for all finite q and in the limit as $q \to \infty$ in terms of the fundamental features of $\bar{\mathbf{A}}$.

Theorem A.1 *For a primitive matrix $\bar{\mathbf{A}}$, all positive integers q, and all $i \in \{1, \ldots, N\}$,*

$$\begin{aligned}\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) &= \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \\ &\quad + \sum_{j=1}^{N}\left[\eta_{ij,q} + \sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q\right][\mathbf{b}]_j \\ &\quad + \mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right),\end{aligned}$$

where $\eta_{ij,q}$, $\left\{\beta_{ij,m_2-s}\right\}_{s=0}^{m_2-1}$ are constants (identified in the accompanying proof), m_2 is the algebraic multiplicity of λ_2, m_3 is the algebraic multiplicity of λ_3, and $\eta_{ij,q} = 0$ whenever $q > N-2$. When $m_2 = 1$,

$$\begin{aligned}\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) &= \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \sum_{j=1}^{N}\left[\eta_{ij,q} + \lambda_2^q\beta_{ij,m_2-0}\right][\mathbf{b}]_j \\ &\quad + \mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right).\end{aligned}$$

The rate of convergence is

$$\left|\frac{\widehat{f}_i^{(q+1)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) - \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)}{\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) - \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)}\right| = \mathcal{O}\left(\left(1+\frac{m_2}{q}\right)^{m_2-1}|\lambda_2|\right).$$

As $q \to \infty$,

$$\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \mathcal{O}\left(q^{m_2-1}|\lambda_2|^q\right).$$

Quantities $\eta_{ij,q}$ and $\{\beta_{ij,m_2-s}\}_{s=0}^{m_2-1}$ are computed by partial fraction decomposition of the ij^{th} element of the resolvent of $\bar{\mathbf{A}}$:

$$[\mathbf{R}(z)]_{ij} = \left[\frac{\text{Adj}(\mathbf{I}-z\bar{\mathbf{A}})}{\det(\mathbf{I}-z\bar{\mathbf{A}})}\right]_{ij} = \eta_{ij}(z) + \frac{\gamma_{ij}}{1-z} + \sum_{s=0}^{m_2-1} \frac{\beta_{ij,m_2-s}}{(1-z\lambda_2)^{m_2-s}} + \sum_{s=0}^{m_3-1} \frac{\beta_{ij,m_3-s}}{(1-z\lambda_3)^{m_3-s}} + \cdots .$$

Eigenvalues λ_j are ordered by decreasing modulus, so that $|\lambda_j| \geq |\lambda_{j'}|$ whenever $j \leq j'$. Adj $(\mathbf{I}-z\bar{\mathbf{A}})$ is the adjugate matrix of $\mathbf{I}-z\bar{\mathbf{A}}$, or the transpose of the cofactor matrix of $\mathbf{I}-z\bar{\mathbf{A}}$, $\eta_{ij}(z) = \eta_{ij,0} + \eta_{ij,1}z + \cdots + \eta_{ij,N-2}z^{N-2}$ is a polynomial of degree at most $N-2$, with $\eta_{ij,q} = 0$ for $q > N-2$, and $\gamma_{ij} = [\mathbf{w}_\infty^T]_j$.

The rate of convergence of $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to its limiting consensus value depends on the second largest eigenvalue modulus, $|\lambda_2|$, the algebraic multiplicity for the second largest eigenvalue, m_2, and the number of iterations, q. The higher the value $|\lambda_2|$, the higher the algebraic multiplicity m_2 for the second largest eigenvalue, and/or the lower the number of iterations q, the slower the rate of convergence of $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ to the consensus value. Theorem A.1 characterizes and accordingly provides insight into finite-round DeGroot learning.

We can additionally characterize $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}}) = [\bar{\mathbf{A}}^q]_{i*}$ for all finite q and in the limit as $q \to \infty$ in terms of the fundamental features of $\bar{\mathbf{A}}$. Since $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]^T \mathbf{b}(N,n)$, following Theorem A.1,

$$\sum_{j=1}^{N} \left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]_j [\mathbf{b}]_j = \sum_{j=1}^{N} \left[[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j + \eta_{ij,q} + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \binom{m_2-s+q-1}{q} \lambda_2^q\right] [\mathbf{b}]_j + \mathcal{O}\left((q+m_3-1)^{m_3-1} |\lambda_3|^q\right).$$

We observe the dependence of $\left[\mathbf{w}_{a,i}^{(q)}\left(\bar{\mathbf{A}}\right)\right]_j$ on matrix primitives $\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]_j$, $\eta_{ij,q}$, and $\left\{\beta_{ij,m_2-s}\right\}_{s=0}^{m_2-1}$ for all $j \in \{1,\ldots,N\}$.

A.4 Section 1.6 Supplemental Material

We first demonstrate how the higher-order features of the distribution of agent weights shape the higher-order features of $G_{X(\bar{\mathbf{A}},N,n)}(t)$ by studying the asymptotic expansion, $J\left(\bar{\mathbf{A}},N,n,t\right)$. To illustrate how skewness of $W\left(\bar{\mathbf{A}}\right)$ generates skewness of $X\left(\bar{\mathbf{A}},N,n\right)$, take the derivative of $J\left(\bar{\mathbf{A}},N,n,t\right)$ from Theorem 1.13 with respect to t to find an approximating probability density function to $g_{\frac{X(\bar{\mathbf{A}},N,n)-EX(\bar{\mathbf{A}},N,n)}{\left(\operatorname{Var}X(\bar{\mathbf{A}},N,n)\right)^{1/2}}}(t)$:[7]

$$J'\left(\bar{\mathbf{A}},N,n,t\right) \equiv \frac{\partial J\left(\bar{\mathbf{A}},N,n,t\right)}{\partial t} = \phi(t) + H_3(t)\,\phi(t)\,C_1\sum_{i=1}^{N}\widehat{w}_i^3$$
$$+ H_4(t)\,\phi(t)\left[C_2\left(\sum_{i=1}^{N}\widehat{w}_i^4 - \frac{3}{N}\right) - \frac{1}{4N}\right] + H_6(t)\,\phi(t)\,C_3\left(\sum_{i=1}^{N}\widehat{w}_i^3\right)^2.$$

The second and fourth terms in the expansion are as follows:

$$H_3(t)\,\phi(t)\,C_1\sum_{i=1}^{N}\widehat{w}_i^3 = -\frac{1-2f}{6\left(2\pi N\right)^{1/2}\left(f\left(1-f\right)\right)^{1/2}} \times \left(3t - t^3\right)e^{-t^2/2}\operatorname{Skew}W\left(\bar{\mathbf{A}}\right),$$

and

$$H_6(t)\,\phi(t)\,C_3\left(\sum_{i=1}^{N}\widehat{w}_i^3\right)^2 = \frac{\left(1-2f\right)^2}{72\left(2\pi\right)^{1/2}Nf(1-f)}$$
$$\times\left(-15 + 45t^2 - 15t^4 + t^6\right)e^{-t^2/2}\left(\operatorname{Skew}W\left(\bar{\mathbf{A}}\right)\right)^2,$$

where $f = \frac{n}{N}$. $\left(3t - t^3\right)e^{-t^2/2}$ is an odd function. Provided that $f < 0.5$ and $\operatorname{Skew}W\left(\bar{\mathbf{A}}\right) > 0$, the second term reallocates mass away from the normal density

[7] The distance, $\left|g_{\frac{X(\bar{\mathbf{A}},N,n)-EX(\bar{\mathbf{A}},N,n)}{\left(\operatorname{Var}X(\bar{\mathbf{A}},N,n)\right)^{1/2}}}(t) - J'\left(\bar{\mathbf{A}},N,n,t\right)\right|$, can also be bounded from above. See line (14) of Robinson (1978), for example.

function $\phi(t)$ to generate positive skewness. If $f > 0.5$ and Skew $W(\bar{\mathbf{A}}) > 0$, the second term reallocates mass away from the normal density function $\phi(t)$ to generate negative skewness. The more heavily skewed the set of agent weights, the more heavily skewed $X(\bar{\mathbf{A}}, N, n)$. As n increases from 1 to $0.5N$, the magnitude of skewness declines, and as n increases from $0.5N$ to $N-1$, the skewness of the distribution changes signs and it increases in magnitude. $\left(-15+45t^2-15t^4+t^6\right)e^{-t^2/2}$ is an even function, so the reallocation of mass away from the normal density function has no effect on the skewness of the distribution. Meanwhile, the relationship between kurtosis of the set of agent weights and kurtosis of $X(\bar{\mathbf{A}}, N, n)$ is a bit more complicated. Heavy-tailedness in the distribution of agent weights can induce heavy-tailedness in $X(\bar{\mathbf{A}}, N, n)$.

We next present some additional results concerning the behavior of $g_{X(\bar{\mathbf{A}},N,n)}(t)$ and more general properties of $X(\bar{\mathbf{A}}, N, n)$:

Theorem A.2 *If $g_{W(\bar{\mathbf{A}})}(t)$ is symmetric, then $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is always symmetric. For any distribution $g_{W(\bar{\mathbf{A}})}(t)$, $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is always symmetric when $f = 0.5$. For all triplets $(\bar{\mathbf{A}}, N, n)$, $g_{X(\bar{\mathbf{A}},N,n)}(t) = g_{X(\bar{\mathbf{A}},N,N-n)}(1-t)$, so* $\operatorname{Var} X(\bar{\mathbf{A}}, N, n) = \operatorname{Var} X(\bar{\mathbf{A}}, N, N-n)$, $\operatorname{Skew} X(\bar{\mathbf{A}}, N, n) = -\operatorname{Skew} X(\bar{\mathbf{A}}, N, N-n)$, *and* $\operatorname{Kurt} X(\bar{\mathbf{A}}, N, n) = \operatorname{Kurt} X(\bar{\mathbf{A}}, N, N-n)$.

Theorem A.2 shows that symmetry in $g_{X(\bar{\mathbf{A}},N,n)}(t)$ arises from symmetry in the distribution of agent weights, and when $f = 0.5$, $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is always symmetric. Theorem A.2 also shows how the second, third, and fourth central moments of the distribution exactly compare when $f = \frac{n}{N}$ and when $f = 1 - \frac{n}{N}$. Theorem A.3 describes properties of the support of the distribution for $X(\bar{\mathbf{A}}, N, n)$ as $f = \frac{n}{N}$ varies:

Theorem A.3 *For $f \in \left[0, \frac{1}{2}\right]$, the width of* $\operatorname{supp} X(\bar{\mathbf{A}}, N, n)$ *weakly increases at a weakly* *decreasing rate, and for* $f \in \left[\frac{1}{2}, 1\right]$*, the width of* $\operatorname{supp} X(\bar{\mathbf{A}}, N, n)$ *weakly decreases at a weakly increasing rate.*

A.5 Statistical Features of the Multivariate Distribution

For every $q \in \mathbb{Z}_{++}$, we study the population vector, $\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, of weighted local relative frequencies of the attribute. Define $\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}, N, n)$ to be the multivariate random variable with realization

$$\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left(\widehat{f}_1^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \cdots \widehat{f}_N^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right)^T$$

and multivariate CDF:

$$G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}, N, n)}(\mathbf{t}) = \frac{1}{|\mathcal{B}(N, n)|} \sum_{\mathbf{b}(N,n) \in \mathcal{B}(N,n)} \mathbb{1}_{\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \leq \mathbf{t}'}$$

where $\mathbf{t} = (t_1 \cdots t_N)^T$, $\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) \leq \mathbf{t}$ holds element-wise, and every configuration is equally likely to occur. We are interested in characterizing $E\,\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}, N, n)$ and $\mathbf{\Sigma}^{(q)}(\bar{\mathbf{A}}, N, n)$, the covariance matrix for $\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}, N, n)$.

To determine the covariance structure of $\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}, N, n)$, we first study the covariance of two scalar random variables, $X_1(\bar{\mathbf{A}}, N, n)$ and $X_2(\bar{\mathbf{A}}, N, n)$. $X_i(\bar{\mathbf{A}}, N, n)$ has realization $x_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = [\mathbf{w}_i(\bar{\mathbf{A}})]^T \mathbf{b}(N, n)$ for $i \in \{1, 2\}$. The $N \times 1$ weighting vectors $\mathbf{w}_1(\bar{\mathbf{A}})$ and $\mathbf{w}_2(\bar{\mathbf{A}})$ may assign different weights to the same agent in the population. For $i \in \{1, 2\}$, random variable $W_i(\bar{\mathbf{A}})$ has realization $[\mathbf{w}_i(\bar{\mathbf{A}})]_j$.

Theorem A.4

$$\operatorname{Cov}(X_1(\bar{\mathbf{A}}, N, n), X_2(\bar{\mathbf{A}}, N, n)) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Cov}(W_1(\bar{\mathbf{A}}), W_2(\bar{\mathbf{A}}))\right).$$

$\mathrm{Cov}\left(W_1\left(\bar{\mathbf{A}}\right), W_2\left(\bar{\mathbf{A}}\right)\right) = \frac{1}{N}\sum_{j=1}^{N}\left(\left[\mathbf{w}_1\left(\bar{\mathbf{A}}\right)\right]_j - \frac{1}{N}\right)\left(\left[\mathbf{w}_2\left(\bar{\mathbf{A}}\right)\right]_j - \frac{1}{N}\right)$. If weights $\left[\mathbf{w}_1\left(\bar{\mathbf{A}}\right)\right]_j$ and $\left[\mathbf{w}_2\left(\bar{\mathbf{A}}\right)\right]_j$ assigned to each agent j strongly covary across agents, then $x_1\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ and $x_2\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ also strongly covary across configurations.

We now characterize the first two moments of $\widehat{\mathbf{F}}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)$ for every $q \in \mathbb{Z}_{++}$:

Theorem A.5 *For all $q \in \mathbb{Z}_{++}$ and for a primitive matrix $\bar{\mathbf{A}}$, $E\,\widehat{\mathbf{F}}^{(q)}\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}\mathbf{1}$. Defining $\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)$ as the covariance matrix for $\widehat{\mathbf{F}}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)$,*

$$\left[\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)\right]_{ik} = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N\,\mathrm{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right)\right),$$

where

$$\mathrm{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right) = \frac{1}{N}\sum_{j=1}^{N}\left(\left[\mathbf{w}_{i,a}^{(q)}\right]_j - \frac{1}{N}\right)\left(\left[\mathbf{w}_{k,a}^{(q)}\right]_j - \frac{1}{N}\right)$$

$$= \frac{1}{N}\sum_{j=1}^{N}\left(\eta_{ij,q} + \sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q + \mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right) + \left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\right)$$
$$\times\left(\eta_{kj,q} + \sum_{s=0}^{m_2-1}\beta_{kj,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q + \mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right) + \left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\right).$$

As $q \to \infty$,

$$\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}}, N, n\right) = \left[\frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}N\right] \times \left[\mathcal{O}\left(q^{2m_2-2}|\lambda_2|^{2q}\right) + \mathrm{Var}\left(W_\infty\left(\bar{\mathbf{A}}\right)\right)\right]\mathbf{1}\mathbf{1}^T,$$

so $\lim_{q\to\infty}\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}}, N, n\right) = \left[\frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}N\,\mathrm{Var}\,W_\infty\left(\bar{\mathbf{A}}\right)\right]\mathbf{1}\mathbf{1}^T$.

The first moment of multivariate random variable $\widehat{\mathbf{F}}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)$ is $\frac{n}{N}\mathbf{1}$: $E\widehat{\mathbf{F}} = \frac{n}{N}\mathbf{1}$ and $E\widehat{\mathbf{F}}^{(q)} = \frac{n}{N}\mathbf{1}$ for all iterations $q > 1$. Along every dimension, the mean equals the global frequency of the attribute. For the second moment, the ik^{th} element of the $N \times N$ covariance matrix $\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)$ directly depends on

$\mathrm{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right)$. The realizations of these random variables $W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right)$ and $W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)$ are respectively the elements in $\mathbf{w}_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right)$ and $\mathbf{w}_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)$, with $\left[\mathbf{w}_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right]^T = \left[\bar{\mathbf{A}}^q\right]_{i*}$ and $\left[\mathbf{w}_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right]^T = \left[\bar{\mathbf{A}}^q\right]_{k*}$.
$\lim_{q\to\infty} \mathrm{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right) = \mathrm{Cov}\left(W_\infty\left(\bar{\mathbf{A}}\right), W_\infty\left(\bar{\mathbf{A}}\right)\right) = \mathrm{Var}\left(W_\infty\left(\bar{\mathbf{A}}\right)\right)$ since $\lim_{q\to\infty} \bar{\mathbf{A}}^q = \mathbf{1}\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T$. As $q \to \infty$, every element in the limiting covariance matrix $\mathbf{\Sigma}^{(\infty)}\left(\bar{\mathbf{A}}, N, n\right)$ therefore has the exact same value, and the corresponding correlation matrix as $q \to \infty$ equals $\mathbf{1}\mathbf{1}^T$.

A.6 Network Topologies that Maximize the Variance of the Distribution

We identify those vectors of agent weights and corresponding network topologies for which the variance of the distribution of possible local relative frequencies of the attribute is maximal. When each configuration of the attribute is equally likely to occur, maximizing the variance of the local relative frequency of the attribute is equivalent to maximizing the variance of the associated set of agent weights:

$$\mathrm{Var}\, X\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N\,\mathrm{Var}\, W\left(\bar{\mathbf{A}}\right)\right).$$

The next theorem specifies the properties of $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ for $\mathrm{Var}\, X\left(\bar{\mathbf{A}}, N, n\right)$ to be maximal:

Theorem A.6 *For a general vector of weights* $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ *with* $0 \le \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i \le 1$, $\forall i \in \{1, \ldots, N\}$, *and* $\mathbf{1}^T\mathbf{w}\left(\bar{\mathbf{A}}\right) = 1$, $\mathrm{Var}\, X\left(\bar{\mathbf{A}}, N, n\right)$ *is maximal when* $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i = 1$ *and* $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j = 0\ \forall j \ne i$.

When agent weights are required to sum to 1, maximizing the variance of $W\left(\bar{\mathbf{A}}\right)$ is equivalent to maximizing the sum of squared agent weights, $\sum_{i=1}^N \left(\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i\right)^2$, with each weight bounded between 0 and 1. For any allocation of weights with

$[\mathbf{w}(\bar{\mathbf{A}})]_i \in [0,1)$ for each $i \in \{1, \ldots, N\}$ and $\mathbf{1}^T\mathbf{w}(\bar{\mathbf{A}}) = 1$, transferring ϵ-mass to the agent weight with the weakly highest value strictly increases the sum of squared weights. It then follows that one agent will have weight 1 while all other agents will have weight 0.

Corollaries A.1 - A.4 build on Theorem A.6 by specifying the necessary values for agents' weights in vectors $\mathbf{w}_{a,i}(\bar{\mathbf{A}})$, $\mathbf{d}_w^-(\bar{\mathbf{A}})$, and $\mathbf{w}_\infty(\bar{\mathbf{A}})$ so that the respective variances of $\widehat{F}_{a,i}(\bar{\mathbf{A}}, N, n)$, $\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$, and $\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ are maximal:

Corollary A.1 (to Theorem A.6) *For the vector of weights* $\mathbf{w}_{a,i}(\bar{\mathbf{A}})$, $\mathrm{Var}\,\widehat{F}_{a,i}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_j = 1$ *and* $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_k = 0$, $\forall k \neq j$. *If the graph features self-loops for every node,* $\mathrm{Var}\,\widehat{F}_{a,i}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_j = 1 - \epsilon$, $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_i = \epsilon$, *and* $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_k = 0$, $\forall k \neq j$, *with* $i \neq j$ *and* $\epsilon > 0$ *small. If* $i = j$, *then* $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_i = 1$ *and* $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]_k = 0$, $\forall k \neq i$.

Corollary A.2 (to Theorem A.6) (1) *Consider* $\mathcal{G}(\mathbf{A})$ *undirected with no self-loops.* $\mathrm{Var}\,\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_j = 1 - \frac{1}{N}$ *and* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_k = \frac{1}{N}\left(\frac{1}{N-1}\right)$, $\forall k \neq j$.

(2) *Consider* $\mathcal{G}(\mathbf{A})$ *undirected with self-loops for every node.* $\mathrm{Var}\,\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_j = 1 - \epsilon$ *and* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_k = \frac{\epsilon}{N-1}$, $\forall k \neq j$ *and* $\epsilon > 0$ *small.*

(3) *Consider* $\mathcal{G}(\mathbf{A})$ *directed with no self-loops except for node* j. $\mathrm{Var}\,\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_j = 1$ *and* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_k = 0$, $\forall k \neq j$.

(4) *Consider* $\mathcal{G}(\mathbf{A})$ *directed with self-loops for every node.* $\mathrm{Var}\,\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_j = 1 - \frac{N-1}{N}\epsilon$ *and* $[\mathbf{d}_w^-(\bar{\mathbf{A}})]_k = \frac{\epsilon}{N}$, $\forall k \neq j$ *and* $\epsilon > 0$ *small.*

For each case, the maximum allowable weight is accorded to a single agent, with the remaining weight accorded to other agents, as needed, to satisfy the properties of $\mathcal{G}(\mathbf{A})$. Figure A.6 plots the extremal graphs $\mathcal{G}(\mathbf{A})$ and their corresponding row-stochastic weighted adjacency matrices $\bar{\mathbf{A}}$ for all four cases in Corollary A.2, setting $j = 15$.

Corollary A.3 (to Theorem A.6) *Consider $\mathcal{G}(\mathbf{A})$ undirected, connected, and aperiodic, with symmetric edge weights for every node. If every node has a self-loop,* $\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j = \frac{N}{3N-2}$ *and* $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_k = \frac{2}{3N-2}$, $\forall k \neq j$.

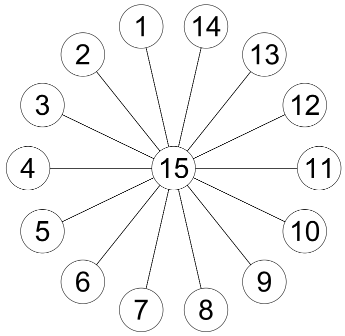
When $\mathcal{G}(\mathbf{A})$ is undirected, connected, and aperiodic, with symmetric edge weights for every node, $\mathbf{w}_\infty(\bar{\mathbf{A}}) = \frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}}$. The maximal degree for a node is N, due to the self-loop, and the minimal degree for a node is 2, since the graph is connected and every node has a self-loop; Corollary A.3 then follows.

Corollary A.4 (to Theorem A.6) *Consider $\mathcal{G}(\mathbf{A})$ directed, Eulerian, and aperiodic, with symmetric edge weights for every node. If every node has a self-loop,* $\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ *is maximal when* $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j = \frac{N}{3N-2}$ *and* $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_k = \frac{2}{3N-2}$ $\forall k \neq j$.

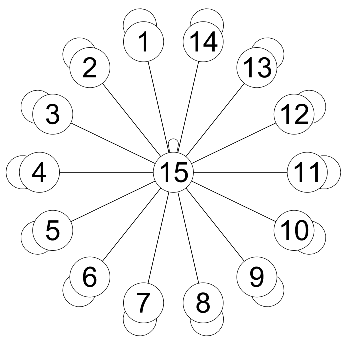
When $\mathcal{G}(\mathbf{A})$ is directed, Eulerian, and aperiodic, with symmetric edge weights for every node, $\mathbf{w}_\infty(\bar{\mathbf{A}}) = \frac{\mathbf{d}^+}{\mathbf{1}^T\mathbf{d}^+}$. The maximal out-degree for a node is N, due to the self-loop, and the minimal out-degree for a node is 2, since the graph is strongly connected and every node has a self-loop.

This last theorem characterizes those networks for which $\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ is maximal given that the underlying graph $\mathcal{G}(\mathbf{A})$ is undirected, aperiodic, and connected with N nodes and M non-self-loop edges, and agents assign equal weight to each of their edges:

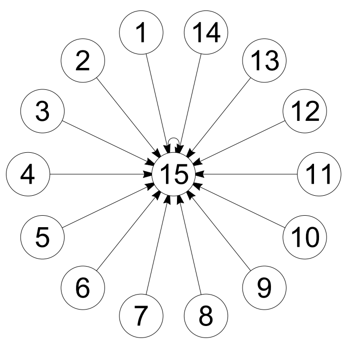
Theorem A.7 *Consider the set of all undirected, aperiodic, connected graphs with N nodes, M non-self-loop edges, symmetric edge weights, and either a self-loop for every node or no self-loop for every node. Either the quasi-star graph $QS(N, M)$ or the quasi-complete graph $QC(N, M)$, with or without self-loops for every node, attains the maximal*

(1) 

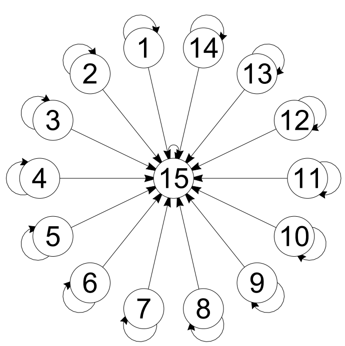
$$\bar{\mathbf{A}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \frac{1}{N-1} & \frac{1}{N-1} & \frac{1}{N-1} & \cdots & 0 \end{pmatrix}$$

(2) 

$$\bar{\mathbf{A}} = \begin{pmatrix} \epsilon & 0 & 0 & \cdots & 0 & 1-\epsilon \\ 0 & \epsilon & 0 & \cdots & 0 & 1-\epsilon \\ 0 & 0 & \epsilon & \cdots & 0 & 1-\epsilon \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \epsilon & 1-\epsilon \\ \frac{\epsilon}{N-1} & \frac{\epsilon}{N-1} & \frac{\epsilon}{N-1} & \cdots & \frac{\epsilon}{N-1} & 1-\epsilon \end{pmatrix}$$

(3) 

$$\bar{\mathbf{A}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

(4) 

$$\bar{\mathbf{A}} = \begin{pmatrix} \epsilon & 0 & 0 & \cdots & 0 & 1-\epsilon \\ 0 & \epsilon & 0 & \cdots & 0 & 1-\epsilon \\ 0 & 0 & \epsilon & \cdots & 0 & 1-\epsilon \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \epsilon & 1-\epsilon \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Figure A.6: *Networks $\mathcal{G}(\mathbf{A})$ and corresponding row-stochastic weighted adjacency matrices $\bar{\mathbf{A}}$ for cases (1)-(4) in Corollary A.2, setting $j = 15$.*

$\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$. *To determine which one maximizes* $\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$*, select the connected graph* $QS(N, M)$ *or* $QC(N, M)$ *with the larger sum of squared degrees arising from non-self-loop edges.*

For the family of undirected, aperiodic, connected graphs $\mathcal{G}(\mathbf{A})$ whose agents assign equal weight to each of their edges, $\mathbf{w}_\infty(\bar{\mathbf{A}}) = \frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}}$ has a closed form. Maximizing $\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ is equivalent to maximizing $\operatorname{Var} W_\infty(\bar{\mathbf{A}})$, and when the number of non-self-loop edges is fixed at M, maximizing $\operatorname{Var} W_\infty(\bar{\mathbf{A}})$ is equivalent to maximizing the sum of squared degrees arising from non-self-loop edges for all nodes in the graph.

Maximizing the sum of squared degrees for simple graphs is a problem studied in graph theory; it is equivalent to maximizing the number of adjacent pairs of edges given that the simple graph has N nodes and M edges.[8] There exist threshold graphs that maximize the graph's sum of squared degrees, of which the quasi-complete graph, $QC(N, M)$, and the quasi-star graph, $QS(N, M)$, are two classes of graphs.[9] The sum of squared degrees for $QC(N, M)$ and $QS(N, M)$

[8]The sum of squared degrees is also known as the first Zagreb index in mathematical chemistry.

[9]As defined by Ahlswede and Katona (1978), a simple quasi-complete graph with N nodes and M edges is constructed as follows: Nodes i and j are connected for all $i, j \leq k$, $i \neq j$, and node $k+1$ is connected to nodes $1, 2, \ldots, \ell$, with k and ℓ uniquely determined by

$$M = \binom{k}{2} + \ell, \quad 0 \leq \ell < k.$$

Conceptually, $QC(N, M)$ consists of the largest possible complete subgraph of nodes $1, \ldots, k$, with edges then added from node $k+1$ to nodes $1, 2, \ldots, \ell$ until the graph has M total edges; any nodes with indices higher than $k+1$ are isolates. If the largest possible complete subgraph of nodes $1, \ldots, k$ has M total edges, then node $k+1$ is also an isolate. Meanwhile, a simple quasi-star graph with N nodes and M edges is constructed as follows: Connect the first $N-k-1$ nodes with every other node and connect node $N-k$ with the first $N-\ell$ nodes, with k and ℓ uniquely determined by

$$\binom{N}{2} - M = \binom{k}{2} + \ell, \quad 0 \leq \ell < k.$$

Conceptually, $QS(N, M)$ consists of assigning the first $N-k-1$ nodes the maximal degree and then can evaluate to the same quantity, so the graph that maximizes the variance of $\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ given M non-self-loop edges is not necessarily unique.

In certain settings, maximizing variance for the distribution of possible local relative frequencies of the attribute is equivalent to maximizing the variance of all possible outcomes of the economy for a fixed aggregate feature:

Example A.2 (Maximizing Variance of the Aggregate Action) *Consider a population of N agents with aggregate action:*

$$a_{agg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \sum_{i=1}^{N} a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \sum_{i=1}^{N}\left[\alpha_i \widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \beta_i\right],$$

with $\alpha_i, \beta_i \in \mathbb{R}$. Let $A_i(\bar{\mathbf{A}}, N, n)$ be a random variable with realization $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$. Then, the variance of the aggregate action, $\operatorname{Var}\left(\sum_{i=1}^{N} A_i(\bar{\mathbf{A}}, N, n)\right)$, is maximal if and only if the variance of the consensus frequency of the attribute, $\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$, is maximal.

To see this equivalence, note that

$$\operatorname{Var}\left(\sum_{i=1}^{N} A_i(\bar{\mathbf{A}}, N, n)\right) = \sum_{i=1}^{N} \operatorname{Var} A_i(\bar{\mathbf{A}}, N, n) + \sum_{i=1}^{N}\sum_{j=1}^{N} \operatorname{Cov}\left(A_i(\bar{\mathbf{A}}, N, n), A_j(\bar{\mathbf{A}}, N, n)\right).$$

With,

$$\begin{aligned}\operatorname{Var} A_i(\bar{\mathbf{A}}, N, n) &= \frac{1}{|\mathcal{B}(N, n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \left(a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) - EA_i(\bar{\mathbf{A}}, N, n)\right)^2 \\ &= \alpha_i^2 \operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n), \text{ and}\end{aligned}$$

---

adding edges from node $N-k$ to nodes $(N-k+1), (N-k+2), \ldots, (N-\ell)$ until the graph has M total edges.

$$\text{Cov}\left(A_i\left(\bar{\mathbf{A}}, N, n\right), A_j\left(\bar{\mathbf{A}}, N, n\right)\right) = \frac{1}{|\mathcal{B}(N,n)|}$$

$$\times \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \left(a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) - EA_i\left(\bar{\mathbf{A}}, N, n\right)\right)\left(a_j\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) - EA_j\left(\bar{\mathbf{A}}, N, n\right)\right)$$

$$= \alpha_i\alpha_j \,\text{Var}\, \widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n\right), \text{ it follows that}$$

$$\text{Var}\left(\sum_{i=1}^{N} A_i\left(\bar{\mathbf{A}}, N, n\right)\right) = \left(\text{Var}\, \widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n\right)\right)\left[\sum_{i=1}^{N}\alpha_i^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j\right].$$

A.7 Multiplicity Results

We characterize those vectors of agent weights and those matrices, $\bar{\mathbf{A}}$, that generate identical distributions of possible local relative frequencies of the attribute and therefore identical distributions of outcomes for the economy. Here, we assume that each configuration of the binary-valued attribute is equally likely. There can potentially be many vectors of agent weights that generate the same distribution $G_{X(\bar{\mathbf{A}},N,n)}(t)$. The next theorem captures this multiplicity:

Theorem A.8 *Consider the general weighting vectors* $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ *and* $\mathbf{w}\left(\bar{\mathbf{A}}'\right)$ *respectively corresponding to the matrices* $\bar{\mathbf{A}}$ *and* $\bar{\mathbf{A}}'$*. For all* $t \in \mathbb{R}$ *and* $n \in \{0, \ldots, N\} \subseteq \mathbb{Z}_+$, $G_{X(\bar{\mathbf{A}}',N,n)}(t) = G_{X(\bar{\mathbf{A}},N,n)}(t)$ *if and only if* $\left[\mathbf{w}\left(\bar{\mathbf{A}}'\right)\right]^T = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{R}$*, where* $\mathbf{R}$ *is any* $N \times N$ *permutation matrix.*

If two weighting vectors are relatable by permutation, $W\left(\bar{\mathbf{A}}'\right) = W\left(\bar{\mathbf{A}}\right)$ so $G_{X(\bar{\mathbf{A}}',N,n)}(t) = G_{X(\bar{\mathbf{A}},N,n)}(t)$.

We now proceed to establish conditions on specific network-derived vectors of agent weights so that their corresponding distributions of possible local relative frequencies of the attribute are identical:

Corollary A.5 (to Theorem A.8) *For all matrices* $\bar{\mathbf{A}}$, $\bar{\mathbf{A}}'$ *and for every pair*
$\left(\mathbf{w}\left(\bar{\mathbf{A}}\right), G_{X(\bar{\mathbf{A}},N,n)}(t)\right) \in \left\{\left\{\left(\mathbf{w}_{a,i}\left(\bar{\mathbf{A}}\right), G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(t)\right)\right\}_{i=1}^{N},\right.$

$\left\{\left(\mathbf{w}_{a,i}^{(q)}\left(\bar{\mathbf{A}}\right), G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}},N,n)}(t)\right)\right\}_{i=1}^{N}, \left(\mathbf{d}_w^{-}\left(\bar{\mathbf{A}}\right), G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)\right),$
$\left(\mathbf{d}_w^{-(q)}\left(\bar{\mathbf{A}}\right), G_{\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}},N,n)}(t)\right), \left(\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right), G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(t)\right)\right\},$
$G_{X(\bar{\mathbf{A}}',N,n)}(t) = G_{X(\bar{\mathbf{A}},N,n)}(t)$ if and only if $\left[\mathbf{w}\left(\bar{\mathbf{A}}'\right)\right]^T = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{R}$, where $\mathbf{R}$ is any $N \times N$ permutation matrix.

As before, $G_{X(\bar{\mathbf{A}}',N,n)}(t) = G_{X(\bar{\mathbf{A}},N,n)}(t)$ if and only if the corresponding vectors of agent weights, $\mathbf{w}\left(\bar{\mathbf{A}}'\right)$ and $\mathbf{w}\left(\bar{\mathbf{A}}\right)$, are permutations of each other.

For each type of network-derived vector, we establish the necessary and sufficient conditions on matrices $\bar{\mathbf{A}}$ and $\bar{\mathbf{A}}'$ for generating identical distributions:

Theorem A.9 (1) *For every $i \in \{1, \ldots, N\}$, $G_{\widehat{F}_i(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(t)$ (respectively $G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}},N,n)}(t)$) if and only if there exists a permutation matrix $\mathbf{R}$ such that $\left[\bar{\mathbf{A}}'\right]_{i*} = \left[\bar{\mathbf{A}}\right]_{i*}\mathbf{R}$ (respectively, $\left[\left(\bar{\mathbf{A}}'\right)^q\right]_{i*} = \left[\bar{\mathbf{A}}^q\right]_{i*}\mathbf{R}$).*
(2) *$G_{\widehat{F}_{avg}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ (respectively $G_{\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_{avg}^{(q)}(\bar{\mathbf{A}},N,n)}(t)$) if and only if there exists a permutation matrix $\mathbf{R}$ such that $\mathbf{1}^T\bar{\mathbf{A}}' = \mathbf{1}^T\bar{\mathbf{A}}\mathbf{R}$ (respectively $\mathbf{1}^T\left(\bar{\mathbf{A}}'\right)^q = \mathbf{1}^T\bar{\mathbf{A}}^q\mathbf{R}$).*
(3) *$G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(t)$ if and only if there exists a permutation matrix $\mathbf{R}$ such that $\bar{\mathbf{A}}'$ and $\mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}$ share a unique dominant left eigenpair.*

The next theorem establishes conditions on $\bar{\mathbf{A}}$ and $\bar{\mathbf{A}}'$ for the multivariate distributions $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t})$ and $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ to equal each other:

Theorem A.10 *$G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ (respectively $G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$) if and only if there exists a permutation matrix $\mathbf{R}$ such that $\bar{\mathbf{A}}' = \bar{\mathbf{A}}\mathbf{R}$ (respectively $\left(\bar{\mathbf{A}}'\right)^q = \bar{\mathbf{A}}^q\mathbf{R}$).*

Multivariate distributions $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t})$ and $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ (respectively $G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}',N,n)}(\mathbf{t})$ and $G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$) equal each other if and only if the set of N columns in $\bar{\mathbf{A}}'$ (respectively $\left(\bar{\mathbf{A}}'\right)^q$) equals the set of N columns in $\bar{\mathbf{A}}$ (respectively $\bar{\mathbf{A}}^q$). In Theorem A.10, agent index matters when establishing conditions for identical multivariate distributions.

Now suppose that an agent's index is irrelevant. We can then establish necessary and sufficient restrictions on $\bar{\mathbf{A}}$ and $\bar{\mathbf{A}}'$ for their affiliated multivariate system-level distributions, $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ and $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t})$, to equal each other, ignoring agent index. Define $\widehat{\mathbf{F}}_S(\bar{\mathbf{A}},N,n)$ as the multivariate random variable with realization $\mathbf{S}\,\widehat{\mathbf{f}}(\bar{\mathbf{A}},\mathbf{b},N,n)$ for permutation matrix $\mathbf{S}$, and let $\widehat{\mathbf{F}}_S^{(q)}(\bar{\mathbf{A}},N,n)$ be the multivariate random variable with realization $\mathbf{S}\,\widehat{\mathbf{f}}^{(q)}(\bar{\mathbf{A}},\mathbf{b},N,n)$:

Theorem A.11 $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = G_{\widehat{\mathbf{F}}_S(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ *(respectively* $G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = G_{\widehat{\mathbf{F}}_S^{(q)}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$*) if and only if there exist permutation matrices* $\mathbf{R}$, $\mathbf{S}$ *such that* $\bar{\mathbf{A}}' = \mathbf{S}\bar{\mathbf{A}}\mathbf{R}$ *(respectively* $(\bar{\mathbf{A}}')^q = \mathbf{S}\bar{\mathbf{A}}^q\mathbf{R}$*).*

When $\bar{\mathbf{A}}' = \bar{\mathbf{A}}\mathbf{R}$ for permutation matrix $\mathbf{R}$, the probability that $\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n) \leq \mathbf{t}$ equals the probability that $\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n) \leq \mathbf{t}$ because the number of configurations $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ for which $\bar{\mathbf{A}}'\mathbf{b}(N,n) \leq \mathbf{t}$ equals the number of configurations $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ for which $\bar{\mathbf{A}}\mathbf{b}(N,n) \leq \mathbf{t}$. When $\bar{\mathbf{A}}' = \mathbf{S}\bar{\mathbf{A}}\mathbf{R}$ for permutation matrices $\mathbf{R}, \mathbf{S}$, the number of configurations $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ for which $\bar{\mathbf{A}}'\mathbf{b}(N,n) \leq \mathbf{t}$ equals the number of configurations $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ for which $\bar{\mathbf{A}}\mathbf{b}(N,n) \leq \mathbf{S}^{-1}\mathbf{t}$.

A.8 Sensitivity of the Distribution to Network Perturbation

We proceed to conduct sensitivity analysis. We study how a perturbation to agents' interaction network affects the distribution of possible local relative frequencies of the attribute via its effects on the relevant network-derived vector of agent weights. The next theorem shows how a perturbation to $\bar{\mathbf{A}}$ alters agents' weights:

Theorem A.12 *Consider the perturbation* $\bar{\mathbf{A}}(\epsilon) = \bar{\mathbf{A}} + \epsilon\,\mathbf{E}$ *about the irreducible, row-stochastic matrix* $\bar{\mathbf{A}}$*, in which* $\mathbf{E1} = \mathbf{0}$ *and* ϵ *is a scalar small enough that* $\bar{\mathbf{A}}(\epsilon)$ *has all non-negative entries. Then,*

$$\mathbf{w}_\infty^T(\epsilon) \approx \mathbf{w}_\infty^T + \epsilon \left(\left. \frac{\partial \mathbf{w}_\infty^T(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} \right),$$

where $\left.\frac{\partial \mathbf{w}_\infty^T(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = \mathbf{w}_\infty^T \mathbf{EZ}$*,* $\mathbf{Z} = \left(\mathbf{I} - \bar{\mathbf{A}} + \mathbf{1p}^T\right)^{-1}$*, and* $\mathbf{p}$ *is any* $N \times 1$ *vector such that* $\mathbf{p}^T\mathbf{1} \neq 0$*. Now letting* $\bar{\mathbf{A}}$ *simply be row-stochastic,*

$$\left[\mathbf{d}_w^-(\epsilon)\right]^T = \left[\mathbf{d}_w^-\right]^T + \frac{\epsilon}{N}\left(\mathbf{1}^T\mathbf{E}\right)$$

exactly; and for every $i \in \{1, \ldots, N\}$*,*

$$\left[\mathbf{w}_{a,i}(\epsilon)\right]^T = \left[\mathbf{w}_{a,i}\right]^T + \epsilon\,[\mathbf{E}]_{i*}$$

exactly. Next, consider the perturbation $\bar{\mathbf{A}}^q(\epsilon) = \bar{\mathbf{A}}^q + \epsilon\mathbf{E}$ *about the row-stochastic matrix* $\bar{\mathbf{A}}^q$*, in which* $\mathbf{E1} = \mathbf{0}$ *and* ϵ *is a scalar small enough that* $\bar{\mathbf{A}}^q(\epsilon)$ *has all non-negative entries. Then,*

$$\left[\mathbf{d}_w^{-(q)}(\epsilon)\right]^T = \left[\mathbf{d}_w^{-(q)}\right]^T + \frac{\epsilon}{N}\left(\mathbf{1}^T\mathbf{E}\right)$$

exactly, and for every $i \in \{1, \ldots, N\}$*,*

$$\left[\mathbf{w}_{a,i}^{(q)}(\epsilon)\right]^T = \left[\mathbf{w}_{a,i}^{(q)}\right]^T + \epsilon\,[\mathbf{E}]_{i*}$$

exactly.

The matrix perturbation preserves row-stochasticity. The effect of the perturbation to $\bar{\mathbf{A}}$ (respectively $\bar{\mathbf{A}}^q$) on $\mathbf{d}_w^-(\bar{\mathbf{A}})$ and $\mathbf{w}_{a,i}(\bar{\mathbf{A}})$ (respectively $\mathbf{d}_w^{-(q)}(\bar{\mathbf{A}})$ and $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})$) for all $i \in \{1, \ldots, N\}$ is exact, but its effect on $\mathbf{w}_\infty(\bar{\mathbf{A}})$ is approximate. The perturbation effect $\left.\frac{\partial \mathbf{w}_\infty^T(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0}$ can be computed by applying existing work on perturbation theory for finite Markov chains (Conlisk, 1985) due to the similarity between a Markov chain's stationary distribution and $\mathbf{w}_{\infty}(\bar{\mathbf{A}})$. Matrix $\mathbf{Z}$ is termed a fundamental matrix (Kemeny and Snell, 1960).[10] The next corollary unambiguously signs the change in certain elements of $\mathbf{w}_{\infty}(\bar{\mathbf{A}})$ following perturbation to $\bar{\mathbf{A}}$:

Corollary A.6 (to Theorem A.12) (1) *For a single row i, if* $[\mathbf{E}]_{ij} > 0$ *and* $[\mathbf{E}]_{ik} < 0$, *with* $[\mathbf{E}]_{ij} + [\mathbf{E}]_{ik} = 0$ *and all other entries in* $\mathbf{E}$ *equal to zero, then*

$$\left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{j}}{\partial\epsilon}\right|_{\epsilon=0} > 0 \quad \text{and} \quad \left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{k}}{\partial\epsilon}\right|_{\epsilon=0} < 0.$$

(2) *For a single row i, if* $[\mathbf{E}]_{ij} > 0$ $\left([\mathbf{E}]_{ij} < 0\right)$ *and* $[\mathbf{E}]_{ik} \leq 0$ $([\mathbf{E}]_{ik} \geq 0)$ *for all* $k \neq j$, *with* $\sum_{k=1}^{n}[\mathbf{E}]_{ik} = 0$ *and all other entries in* $\mathbf{E}$ *equal to zero, then*

$$\left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{j}}{\partial\epsilon}\right|_{\epsilon=0} > 0 \qquad \left(\left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{j}}{\partial\epsilon}\right|_{\epsilon=0} < 0\right).$$

If there exists a perturbation in which agent i transfers weight to agent j from agent k, then agent j's weight under consensus unambiguously increases, and agent k's weight under consensus unambiguously decreases. If there exists a separate perturbation in which agent i transfers weight to agent j from all other agents $k \neq j$, then the weight of agent j under consensus unambiguously increases. This second case arises when agent i adds a directed linkage to agent j and equally weights each of his out-edges.[11]

The effects of network perturbation on $G_{X(\bar{\mathbf{A}},N,n)}(t)$ can be characterized as follows:

Theorem A.13 *Consider the perturbation* $\bar{\mathbf{A}}(\epsilon) = \bar{\mathbf{A}} + \epsilon\,\mathbf{E}$ *about row-stochastic matrix*

[10] The first part of Theorem A.12, the effect of a matrix perturbation on $\mathbf{w}_{\infty}(\bar{\mathbf{A}})$, is similar to Theorem 6 in Golub and Jackson (2007), but the method of proof differs.

[11] Corollary A.6 is similar to Corollaries 1-2 in Golub and Jackson (2007), but the method of proof differs.

$\bar{\mathbf{A}}$, *in which* $\mathbf{E1} = \mathbf{0}$ *and* ϵ *is a scalar small enough that* $\bar{\mathbf{A}}(\epsilon)$ *has all non-negative entries. Then,*

$$EX\left(\bar{\mathbf{A}}(\epsilon), N, n\right) = EX\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N},$$
$$\operatorname{Var} X\left(\bar{\mathbf{A}}(\epsilon), N, n\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W(\epsilon)\right), \text{ and}$$

$$J\left(\bar{\mathbf{A}}(\epsilon), N, n, t\right) = \Phi(t) - H_2(t)\,\phi(t)\,C_1 N^{-1/2} \operatorname{Skew} W(\epsilon) - H_3(t)\,\phi(t)\left[C_2\left(N^{-1} \text{Excess Kurtosis}\, W(\epsilon)\right) - \frac{1}{4N}\right] - H_5(t)\,\phi(t)\,C_3 N^{-1}\left(\operatorname{Skew} W(\epsilon)\right)^2,$$

with C_1, C_2, C_3, $\phi(t)$, *and* $H_i(t)\,\phi(t)$ *defined in Theorem 1.13,* $\widehat{w}_i(\epsilon) = \frac{[\mathbf{w}(\epsilon)]_i - EW(\epsilon)}{\sqrt{N \operatorname{Var} W(\epsilon)}}$, *and specific cases of perturbed weights* $\mathbf{w}(\epsilon)$ *listed in Theorem A.12.*

If the perturbed vector of weights is $\mathbf{w}_\infty(\bar{\mathbf{A}})$, we assume that $\bar{\mathbf{A}}$ is primitive; otherwise, this assumption is not necessary. Following perturbation to $\bar{\mathbf{A}}$, the first moment of the distribution remains fixed at $\frac{n}{N}$. If $\operatorname{Var} W(\epsilon) > \operatorname{Var} W(0)$, then the second moment of $X(\bar{\mathbf{A}}, N, n)$ increases after perturbation. Changes to $G_{X(\bar{\mathbf{A}},N,n)}(t)$ depend on how the perturbation to $\bar{\mathbf{A}}$ affects the variance and higher-order moments of the set of agent weights.

A.9 Section 1.7 Examples

Example 1.7 (First Two Moments of $\mathbf{X}\left(\bar{\mathbf{A}}, \mathbf{N}, \mathbf{n}, (\gamma_\mathbf{i})_{\mathbf{i}=\mathbf{1}}^{\mathbf{N}}\right)$, $\Theta = 1$) *Consider an economic system with N agents. When* $\Theta = 1$, *so that* $\phi_i = \rho_1$ *for every agent* $i \in \{1, \ldots, N\}$, $EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = EX\left(\bar{\mathbf{A}}, N, n\right)$ *and* $\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = \operatorname{Var} X\left(\bar{\mathbf{A}}, N, n\right)$, *where* $X\left(\bar{\mathbf{A}}, N, n\right)$ *is the random variable of interest when every configuration* $\mathbf{b}(N, n) \in \mathcal{B}(N, n)$ *is equally likely.*

When $\Theta = 1$, we know that every agent $i \in \{1, \ldots, N\}$ has the same conditional probability that $B_i = 1$, and every configuration $\mathbf{b}(N, n) \in \mathcal{B}(N, n)$ is equally likely. From Theorem 1.17,

$$\sum_{i=1}^{N} [\widehat{\boldsymbol{\mu}}]_1 = n,$$

so $[\widehat{\boldsymbol{\mu}}]_1 = \frac{n}{N}$ and

$$EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = \sum_{i=1}^{N} [\mathbf{w}(\bar{\mathbf{A}})]_i [\boldsymbol{\mu}]_i = \sum_{i=1}^{N} [\mathbf{w}(\bar{\mathbf{A}})]_i [\widehat{\boldsymbol{\mu}}]_1 = \frac{n}{N}.$$

Meanwhile, $[\zeta]_i = [\widehat{\boldsymbol{\mu}}]_1 (1 - [\widehat{\boldsymbol{\mu}}]_1) = \frac{n}{N}\left(1 - \frac{n}{N}\right)$, and $\mathbf{1}^T\zeta = \frac{n}{N}\left(1 - \frac{n}{N}\right) N$. With

$$\begin{aligned}
\boldsymbol{\Sigma} &= \frac{N}{N-1}\left(\operatorname{diag}\zeta - \frac{\zeta\zeta^T}{\mathbf{1}^T\zeta}\right), \\
[\boldsymbol{\Sigma}]_{ij} &= \frac{N}{N-1}\left(-\frac{[\zeta]_i [\zeta]_j}{\mathbf{1}^T\zeta}\right) = [\boldsymbol{\Sigma}]_{ji}, \text{ for } i \neq j, \text{ and} \\
[\boldsymbol{\Sigma}]_{ii} &= \frac{N}{N-1}\left([\zeta]_i - \frac{[\zeta]_i [\zeta]_j}{\mathbf{1}^T\zeta}\right) = \frac{N}{N-1}[\zeta]_i + [\boldsymbol{\Sigma}]_{ij}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = [\mathbf{w}(\bar{\mathbf{A}})]^T \boldsymbol{\Sigma} [\mathbf{w}(\bar{\mathbf{A}})] \\
&= \sum_{i=1}^{N} ([\mathbf{w}(\bar{\mathbf{A}})]_i)^2 [\boldsymbol{\Sigma}]_{ii} + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} [\mathbf{w}(\bar{\mathbf{A}})]_i [\mathbf{w}(\bar{\mathbf{A}})]_j [\boldsymbol{\Sigma}]_{ij} \\
&= \sum_{i=1}^{N} ([\mathbf{w}(\bar{\mathbf{A}})]_i)^2 \left(\frac{N}{N-1}[\zeta]_i + [\boldsymbol{\Sigma}]_{ij}\right) + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} [\mathbf{w}(\bar{\mathbf{A}})]_i [\mathbf{w}(\bar{\mathbf{A}})]_j [\boldsymbol{\Sigma}]_{ij} \\
&= \left[\sum_{i=1}^{N} ([\mathbf{w}(\bar{\mathbf{A}})]_i)^2 \left(\frac{N}{N-1}[\zeta]_i\right)\right] + ([\mathbf{w}(\bar{\mathbf{A}})]_1 + [\mathbf{w}(\bar{\mathbf{A}})]_2 + \cdots + [\mathbf{w}(\bar{\mathbf{A}})]_N)^2 [\boldsymbol{\Sigma}]_{ij} \\
&= \left[\sum_{i=1}^{N} ([\mathbf{w}(\bar{\mathbf{A}})]_i)^2 \left(\frac{N}{N-1}[\zeta]_i\right)\right] + [\boldsymbol{\Sigma}]_{ij} \\
&= \frac{N}{N-1}\frac{n}{N}\left(1 - \frac{n}{N}\right)\left[([\mathbf{w}(\bar{\mathbf{A}})]_1)^2 + ([\mathbf{w}(\bar{\mathbf{A}})]_2)^2 + \cdots + ([\mathbf{w}(\bar{\mathbf{A}})]_N)^2 - \frac{1}{N}\right] \\
&= \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W(\bar{\mathbf{A}})\right)
\end{aligned}$$

because

$$N \operatorname{Var} W(\bar{\mathbf{A}}) = N\left(\frac{1}{N}\sum_{i=1}^{N}\left([\mathbf{w}(\bar{\mathbf{A}})]_i - \frac{1}{N}\right)^2\right) = \left(\sum_{i=1}^{N}\left([\mathbf{w}(\bar{\mathbf{A}})]_i\right)^2\right) - \frac{1}{N}.$$

Therefore, $EX\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right) = \frac{n}{N}$ and $\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W(\bar{\mathbf{A}})\right)$. We have thus recovered the first two moments of $X\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right)$ when every configuration $\mathbf{b}(N,n)$ is uniformly selected from $\mathcal{B}(N,n)$.

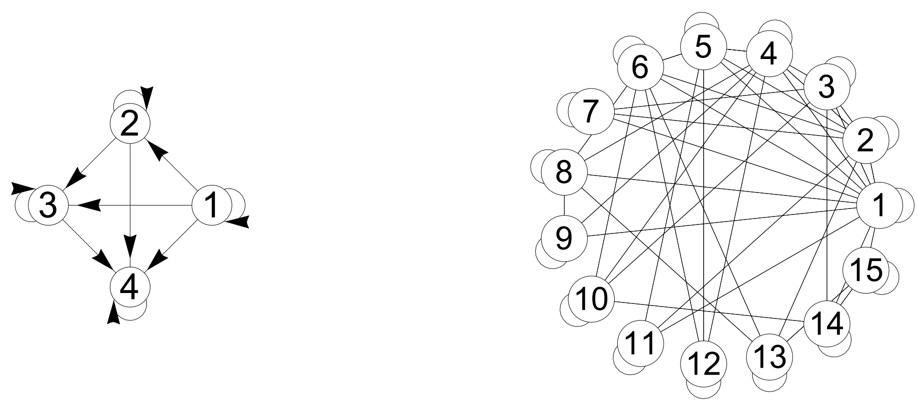


Figure A.7: *Graph $\mathcal{G}(\mathbf{A}_1)$ for Example A.3 (left) and graph $\mathcal{G}(\mathbf{A}_2)$ for Example A.4 (right).*

Example A.3 (First Two Moments of $\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right)$, $\Theta = 2$) *Consider an economy with $N = 4$ agents and underlying network $\mathcal{G}(\mathbf{A}_1)$ featured on the left side of Figure A.7. Assume that agents assign an equal weight to each of their out-edges. There are $\Theta = 2$ categories of agents. In category 1, $\phi_1 = \phi_2 = \phi_3 = \rho_1 = \frac{1}{5}$, and in category 2, $\phi_4 = \rho_2 = \frac{1}{3}$. Given that $f = 0.25$, $E\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^{N}\right) \approx 0.302$ and* Std. Dev. $\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^{N}\right) \approx 0.180$.

We set $k = 2$, and computing the odds ratio, $\widehat{\psi}_1$, we have:

$$\widehat{\psi}_1 = \frac{\frac{\rho_1}{1-\rho_1}}{\frac{\rho_2}{1-\rho_2}} = \frac{\frac{1/5}{1-1/5}}{\frac{1/3}{1-1/3}} = \frac{1}{2}.$$

To compute the first two moments of $\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right)$, we must solve for the following six variables: $[\widehat{\boldsymbol{\mu}}]_1$, $[\widehat{\boldsymbol{\mu}}]_2$, $\left[\widehat{\boldsymbol{\Sigma}}\right]_{11}$, $\left[\widehat{\boldsymbol{\Sigma}}\right]_{22}$, $\left[\widehat{\boldsymbol{\Sigma}}\right]_{12}$, and $\left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1$. $\widehat{\boldsymbol{\Sigma}}^{Cov}$ is a $\Theta \times 1$ vector whose θ^{th} element equals $\mathrm{Cov}\left(B_i, B_j\right)$ for agent i and agent j both in the same category θ. From $\widehat{\boldsymbol{\mu}}$, $\widehat{\boldsymbol{\Sigma}}$, and $\left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1$, we can construct $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, and given $\mathbf{d}_w^-\left(\bar{\mathbf{A}}_1\right)$, we can determine the first two moments of $\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^{N}\right)$.

Following Theorem 1.17, the system of six equations is below:

$$(1) \qquad 3\,[\widehat{\boldsymbol{\mu}}]_1 + [\widehat{\boldsymbol{\mu}}]_2 = n$$

$$(2) \qquad \frac{1}{2} = \frac{[\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right) - \left[\widehat{\boldsymbol{\Sigma}}\right]_{12}}{\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right)[\widehat{\boldsymbol{\mu}}]_2 - \left[\widehat{\boldsymbol{\Sigma}}\right]_{12}}$$

$$(3) \qquad \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} = \frac{N}{N-1}\left([\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right) - \frac{\left([\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right)\right)^2}{3\,[\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right) + [\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right)}\right)$$

$$(4) \qquad \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} = -\frac{N}{N-1}\left(\frac{[\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right)[\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right)}{3\,[\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right) + [\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right)}\right)$$

$$(5) \qquad \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} = \frac{N}{N-1}\left([\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right) - \frac{\left([\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right)\right)^2}{3\,[\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right) + [\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right)}\right)$$

$$(6) \qquad \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 = -\frac{N}{N-1}\left(\frac{\left([\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right)\right)^2}{3\,[\widehat{\boldsymbol{\mu}}]_1\left(1 - [\widehat{\boldsymbol{\mu}}]_1\right) + [\widehat{\boldsymbol{\mu}}]_2\left(1 - [\widehat{\boldsymbol{\mu}}]_2\right)}\right).$$

Then,

$$\boldsymbol{\mu} = \begin{pmatrix} [\widehat{\mu}]_1 \\ [\widehat{\mu}]_1 \\ [\widehat{\mu}]_1 \\ [\widehat{\mu}]_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \\ \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 & \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \\ \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 & \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} \end{pmatrix}.$$

With $[\widehat{\mu}]_1 \geq 0$ and $[\widehat{\mu}]_2 \geq 0$, we find that

$$\begin{aligned} [\widehat{\mu}]_1 &\approx 0.202, \\ [\widehat{\mu}]_2 &\approx 0.395, \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} &\approx 0.167, \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} &\approx 0.213, \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} &\approx -0.0711, \text{ and} \\ \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_1 &\approx -0.0479. \end{aligned}$$

With $\mathbf{d}_w^- \left(\bar{\mathbf{A}}_1\right) \approx \begin{pmatrix} 0.0625 & 0.146 & 0.271 & 0.521 \end{pmatrix}^T$,

$$\begin{aligned} E\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^N\right) &\approx 0.302, \\ \text{Var}\,\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^N\right) &\approx 0.0325, \text{ and} \\ \text{Std. Dev.}\,\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^N\right) &\approx 0.180. \end{aligned}$$

$E\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n, (\gamma_i)_{i=1}^N\right) \approx 0.302 > 0.25 = E\widehat{F}_{avg}\left(\bar{\mathbf{A}}_1, N, n\right)$, the mean of the distribution when $n = 1$ and every configuration is equally likely.

Example A.4 (First Two Moments of $\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right)$, $\Theta = 3$) *Consider an economy with $N = 15$ agents and underlying network $\mathcal{G}(\mathbf{A}_2)$ featured on the right side of Figure A.7. Assume that agents assign an equal weight to each of their edges. There are $\Theta = 3$ categories of agents. In category 1, $\phi_i = \rho_1 = \frac{2}{3}$ for agents $i \in \{1, \ldots, 5\}$; in category 2, $\phi_i = \rho_2 = \frac{2}{5}$ for agents $i \in \{6, \ldots, 10\}$; and in category 3, $\phi_i = \rho_3 = \frac{1}{4}$ for agents $i \in \{11, \ldots, 15\}$. Given that $f = 0.20$, $E\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^{N}\right) \approx 0.238$ and* $\text{Std. Dev.}\, \widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^{N}\right) \approx 0.0406.$

We set $k = 3$, and computing the odds ratios $\widehat{\psi}_1$ and $\widehat{\psi}_2$, we have:

$$\widehat{\psi}_1 = \frac{\frac{\rho_1}{1-\rho_1}}{\frac{\rho_3}{1-\rho_3}} = \frac{\frac{2/3}{1-2/3}}{\frac{1/4}{1-1/4}} = 6 \quad \text{and} \quad \widehat{\psi}_2 = \frac{\frac{\rho_2}{1-\rho_2}}{\frac{\rho_3}{1-\rho_3}} = \frac{\frac{2/5}{1-2/5}}{\frac{1/4}{1-1/4}} = 2.$$

With $\Theta = 3$ categories and five agents in each category, there are $2\Theta + \binom{\Theta}{2} + \sum_{\theta=1}^{\Theta} \mathbb{1}_{s_\theta > 1} = 12$ variables and 12 equations. Following Theorem 1.17, the system of twelve equations is below:

$$(1) \qquad 5\,[\widehat{\mu}]_1 + 5\,[\widehat{\mu}]_2 + 5\,[\widehat{\mu}]_3 = n$$

$$(2) \qquad 6 = \frac{[\widehat{\mu}]_1\,(1-[\widehat{\mu}]_3) - \left[\widehat{\Sigma}\right]_{13}}{(1-[\widehat{\mu}]_1)\,[\widehat{\mu}]_3 - \left[\widehat{\Sigma}\right]_{13}}$$

$$(3) \qquad 2 = \frac{[\widehat{\mu}]_2\,(1-[\widehat{\mu}]_3) - \left[\widehat{\Sigma}\right]_{23}}{(1-[\widehat{\mu}]_2)\,[\widehat{\mu}]_3 - \left[\widehat{\Sigma}\right]_{23}}$$

$$(4) \qquad \left[\widehat{\Sigma}\right]_{11} = \frac{N}{N-1}\left([\widehat{\mu}]_1\,(1-[\widehat{\mu}]_1) - \frac{([\widehat{\mu}]_1\,(1-[\widehat{\mu}]_1))^2}{\mathcal{M}}\right)$$

$$(5) \qquad \left[\widehat{\Sigma}\right]_{22} = \frac{N}{N-1}\left([\widehat{\mu}]_2\,(1-[\widehat{\mu}]_2) - \frac{([\widehat{\mu}]_2\,(1-[\widehat{\mu}]_2))^2}{\mathcal{M}}\right)$$

$$(6) \qquad \left[\widehat{\Sigma}\right]_{33} = \frac{N}{N-1}\left([\widehat{\mu}]_3\,(1-[\widehat{\mu}]_3) - \frac{([\widehat{\mu}]_3\,(1-[\widehat{\mu}]_3))^2}{\mathcal{M}}\right)$$

$$(7) \qquad \left[\widehat{\Sigma}\right]_{12} = -\frac{N}{N-1}\left(\frac{[\widehat{\mu}]_1\,(1-[\widehat{\mu}]_1)\,[\widehat{\mu}]_2\,(1-[\widehat{\mu}]_2)}{\mathcal{M}}\right)$$

$$(8) \qquad \left[\widehat{\Sigma}\right]_{13} = -\frac{N}{N-1}\left(\frac{[\widehat{\mu}]_1\,(1-[\widehat{\mu}]_1)\,[\widehat{\mu}]_3\,(1-[\widehat{\mu}]_3)}{\mathcal{M}}\right)$$

$$(9) \qquad \left[\widehat{\Sigma}\right]_{23} = -\frac{N}{N-1}\left(\frac{[\widehat{\mu}]_2\,(1-[\widehat{\mu}]_2)\,[\widehat{\mu}]_3\,(1-[\widehat{\mu}]_3)}{\mathcal{M}}\right)$$

$$(10) \qquad \left[\widehat{\Sigma}^{Cov}\right]_1 = -\frac{N}{N-1}\left(\frac{([\widehat{\mu}]_1\,(1-[\widehat{\mu}]_1))^2}{\mathcal{M}}\right)$$

$$(11) \qquad \left[\widehat{\Sigma}^{Cov}\right]_2 = -\frac{N}{N-1}\left(\frac{([\widehat{\mu}]_2\,(1-[\widehat{\mu}]_2))^2}{\mathcal{M}}\right)$$

$$(12) \qquad \left[\widehat{\Sigma}^{Cov}\right]_3 = -\frac{N}{N-1}\left(\frac{([\widehat{\mu}]_3\,(1-[\widehat{\mu}]_3))^2}{\mathcal{M}}\right),$$

where $\mathcal{M} = 5\,[\widehat{\mu}]_1\,(1-[\widehat{\mu}]_1) + 5\,[\widehat{\mu}]_2\,(1-[\widehat{\mu}]_2) + 5\,[\widehat{\mu}]_3\,(1-[\widehat{\mu}]_3)$. With $\widehat{\boldsymbol{\mu}} \geq \mathbf{0}$, we find that

$$[\widehat{\mu}]_1 \approx 0.370, \qquad [\widehat{\mu}]_2 \approx 0.151, \qquad [\widehat{\mu}]_3 \approx 0.0791,$$

$$\left[\widehat{\Sigma}\right]_{11} \approx 0.223, \qquad \left[\widehat{\Sigma}\right]_{22} \approx 0.129, \qquad \left[\widehat{\Sigma}\right]_{33} \approx 0.0754,$$

$$\left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \approx -0.0147, \qquad \left[\widehat{\boldsymbol{\Sigma}}\right]_{13} \approx -0.00838, \qquad \left[\widehat{\boldsymbol{\Sigma}}\right]_{23} \approx -0.00461,$$
$$\left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} \approx -0.0268, \qquad \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} \approx -0.00810, \qquad \text{and} \qquad \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{3} \approx -0.00262.$$

From the $\Theta \times 1$ vector $\widehat{\boldsymbol{\mu}}$ and the $\Theta \times \Theta$ matrix $\widehat{\boldsymbol{\Sigma}}$, we can construct the $N \times 1$ vector $\boldsymbol{\mu}$ and the $N \times N$ matrix $\boldsymbol{\Sigma}$. Given the vector of consensus weights $\mathbf{w}_\infty\left(\bar{\mathbf{A}}_2\right)$,

$$E\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^N\right) \approx 0.238,$$
$$\text{Var}\,\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^N\right) \approx 0.00165, \text{ and}$$
$$\text{Std. Dev.}\,\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^N\right) \approx 0.0406.$$

$E\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^N\right) \approx 0.238 > 0.20 = E\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n\right)$, the mean of the distribution when $n = 3$ and every configuration is equally likely.
$\text{Std. Dev.}\,\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n, (\gamma_i)_{i=1}^N\right) \approx 0.0406 > 0.0404 \approx \text{Std. Dev.}\,\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}}_2, N, n\right)$, the standard deviation of the distribution when $n = 3$ and every configuration is equally likely. When agents differ in their probabilities of having the attribute's unit value, the mean of the probability distribution of possible local relative frequencies of the attribute can potentially markedly diverge from the attribute's global relative frequency.

Example 1.4 (Configurations Unequally Likely, $\widehat{\psi}_1 = 9.42$) *Suppose that media outlets engage in "fair and balanced" reporting, providing equal air time (or equal space for hard-copy publications) to those agents who are employed and unemployed. Setting* $\rho_1 = 0.50$ *and* $\rho_2 = 0.096$, $E\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = 0.194$ *and* $\text{Std. Dev.}\,\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right) = 0.00452$.

Following Theorem 1.17, to compute the first two moments of $\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^N\right)$ for a general $\widehat{\psi}_1$, we must solve for seven variables: $[\widehat{\boldsymbol{\mu}}]_1$, $[\widehat{\boldsymbol{\mu}}]_2$,

$\left[\widehat{\mathbf{\Sigma}}\right]_{11}$, $\left[\widehat{\mathbf{\Sigma}}\right]_{22}$, $\left[\widehat{\mathbf{\Sigma}}\right]_{12}$, $\left[\widehat{\mathbf{\Sigma}}^{Cov}\right]_{1}$, and $\left[\widehat{\mathbf{\Sigma}}^{Cov}\right]_{2}$. The system of seven equations is below:

$$
\begin{aligned}
&(1) \qquad x\,[\widehat{\boldsymbol{\mu}}]_1 + (N-x)\,[\widehat{\boldsymbol{\mu}}]_2 = n \\
&(2) \qquad \widehat{\psi}_1 = \frac{[\widehat{\boldsymbol{\mu}}]_1\,(1-[\widehat{\boldsymbol{\mu}}]_2) - \left[\widehat{\mathbf{\Sigma}}\right]_{12}}{(1-[\widehat{\boldsymbol{\mu}}]_1)\,[\widehat{\boldsymbol{\mu}}]_2 - \left[\widehat{\mathbf{\Sigma}}\right]_{12}} \\
&(3) \qquad \left[\widehat{\mathbf{\Sigma}}\right]_{11} = \frac{N}{N-1}\left([\widehat{\boldsymbol{\mu}}]_1\,(1-[\widehat{\boldsymbol{\mu}}]_1) - \frac{([\widehat{\boldsymbol{\mu}}]_1\,(1-[\widehat{\boldsymbol{\mu}}]_1))^2}{\mathcal{M}}\right) \\
&(4) \qquad \left[\widehat{\mathbf{\Sigma}}\right]_{12} = -\frac{N}{N-1}\left(\frac{[\widehat{\boldsymbol{\mu}}]_1\,(1-[\widehat{\boldsymbol{\mu}}]_1)\,[\widehat{\boldsymbol{\mu}}]_2\,(1-[\widehat{\boldsymbol{\mu}}]_2)}{\mathcal{M}}\right) \\
&(5) \qquad \left[\widehat{\mathbf{\Sigma}}\right]_{22} = \frac{N}{N-1}\left([\widehat{\boldsymbol{\mu}}]_2\,(1-[\widehat{\boldsymbol{\mu}}]_2) - \frac{([\widehat{\boldsymbol{\mu}}]_2\,(1-[\widehat{\boldsymbol{\mu}}]_2))^2}{\mathcal{M}}\right) \\
&(6) \qquad \left[\widehat{\mathbf{\Sigma}}^{Cov}\right]_{1} = -\frac{N}{N-1}\left(\frac{([\widehat{\boldsymbol{\mu}}]_1\,(1-[\widehat{\boldsymbol{\mu}}]_1))^2}{\mathcal{M}}\right) \\
&(7) \qquad \left[\widehat{\mathbf{\Sigma}}^{Cov}\right]_{2} = -\frac{N}{N-1}\left(\frac{([\widehat{\boldsymbol{\mu}}]_2\,(1-[\widehat{\boldsymbol{\mu}}]_2))^2}{\mathcal{M}}\right),
\end{aligned}
$$

with $\mathcal{M} = x\,([\widehat{\boldsymbol{\mu}}]_1\,(1-[\widehat{\boldsymbol{\mu}}]_1)) + (N-x)\,([\widehat{\boldsymbol{\mu}}]_2\,(1-[\widehat{\boldsymbol{\mu}}]_2))$. Then,

$$
E\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right) = [\widehat{\boldsymbol{\mu}}]_1\left(\sum_{\substack{i\in\{1,\ldots,N\} \\ \text{s.t. } \phi_i=\rho_1}}\left[\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)\right]_i\right) + [\widehat{\boldsymbol{\mu}}]_2\left(\sum_{\substack{i\in\{1,\ldots,N\} \\ \text{s.t. } \phi_i=\rho_2}}\left[\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)\right]_i\right)
$$

and

$$
\operatorname{Var}\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n, (\gamma_i)_{i=1}^{N}\right) = \left[\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{\Sigma}\left[\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)\right],
$$

with the $N \times N$ matrix

$$\boldsymbol{\Sigma} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ x \\ x+1 \\ x+2 \\ \vdots \\ N \end{array} \begin{array}{c} 1 & 2 & \cdots & x & x+1 & x+2 & \cdots & N \\ \left(\begin{array}{cccccccc} \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \\ \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} & & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \\ \vdots & & \ddots & & \vdots & \vdots & & \vdots \\ \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} & & \left[\widehat{\boldsymbol{\Sigma}}\right]_{11} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} & & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \ddots & \\ \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \cdots & \left[\widehat{\boldsymbol{\Sigma}}\right]_{12} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} & \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} & & \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} \end{array} \right) \end{array}.$$

With $[\widehat{\boldsymbol{\mu}}]_1 \geq 0$ and $[\widehat{\boldsymbol{\mu}}]_2 \geq 0$, we find that

$$[\widehat{\boldsymbol{\mu}}]_1 \approx 0.500, \quad [\widehat{\boldsymbol{\mu}}]_2 \approx 0.0959,$$

$$\left[\widehat{\boldsymbol{\Sigma}}\right]_{11} \approx 0.250, \quad \left[\widehat{\boldsymbol{\Sigma}}\right]_{22} \approx 0.0867,$$

$$\left[\widehat{\boldsymbol{\Sigma}}\right]_{12} \approx -1.82 \times 10^{-9}, \quad \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{1} \approx -5.24 \times 10^{-9}, \quad \text{and} \quad \left[\widehat{\boldsymbol{\Sigma}}^{Cov}\right]_{2} \approx -6.30 \times 10^{-10}.$$

A.10 Section 1.8 Examples

Example A.5 (Distribution of Agent *i*'s Action, Affine Rule) *Suppose that agent i's action is:* $a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \alpha_i \widehat{f_i}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) + \beta_i$*. If each configuration of the attribute is equally likely to occur,* $EA_i\left(\bar{\mathbf{A}}, N, n\right) = \alpha_i \frac{n}{N} + \beta_i$ *and* $\operatorname{Var} A_i\left(\bar{\mathbf{A}}, N, n\right) = \alpha_i^2 \frac{n}{N}\left(1 - \frac{n}{N}\right) \frac{N}{N-1} N \operatorname{Var} W_{a,i}\left(\bar{\mathbf{A}}\right)$*. To compute the upper and lower bounds on the support of* $A_i\left(\bar{\mathbf{A}}, N, n\right)$*, construct the ordered multiset* $\{w_s\}_{s=1}^{N}$ *from the elements of* $\mathbf{w}_{a,i}\left(\bar{\mathbf{A}}\right)$ *so* *that $w_s \leq w_{s'}$ whenever $s \leq s'$. Assuming that $\alpha_i > 0$,*

$$\max \operatorname{supp} A_i\left(\bar{\mathbf{A}}, N, n\right) = \alpha_i \left(\sum_{s=N-n+1}^{N} w_s\right) + \beta_i, \text{ and}$$

$$\min \operatorname{supp} A_i\left(\bar{\mathbf{A}}, N, n\right) = \alpha_i \left(\sum_{s=1}^{n} w_s\right) + \beta_i.$$

Provided that condition (c) of Theorem 1.13 holds,

$$G_{A_i(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{t - \alpha_i E\widehat{F}_i\left(\bar{\mathbf{A}}, N, n\right) - \beta_i}{\alpha_i \left(\operatorname{Var} \widehat{F}_i\left(\bar{\mathbf{A}}, N, n\right)\right)^{1/2}}\right),$$

with $J\left(\cdot\right)$ defined in Theorem 1.13 and $\widehat{w}_j = \frac{\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_j - EW_{a,i}(\bar{\mathbf{A}})}{\left(N \operatorname{Var} W_{a,i}(\bar{\mathbf{A}})\right)^{1/2}} = \frac{\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_j - \frac{1}{N}}{\left(\sum_{k=1}^{N}\left(\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_k - \frac{1}{N}\right)^2\right)^{1/2}}.$

When agent i's action linearly depends on $\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$, we can solve for $G_{A_i(\bar{\mathbf{A}},N,n)}\left(t\right)$ for every feasible population size, network topology, and prevalence of the attribute in the population.

Example A.6 (Distribution of Agent *i*'s Action, Nonlinear Rule) *Suppose that agent i's action is: $a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \alpha_i \log \widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) + \beta_i$. Assume that $n \geq 1$ and $\mathbf{w}_{a,i}^{(q)}\left(\bar{\mathbf{A}}\right) > \mathbf{0}$ element-wise, so that $\log \widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) > -\infty$. To compute the upper and lower bounds of $A_i\left(\bar{\mathbf{A}}, N, n\right)$, construct the ordered multiset $\{w_s\}_{s=1}^{N}$ from the elements of $\mathbf{w}_{a,i}^{(q)}\left(\bar{\mathbf{A}}\right)$ so that $w_s \leq w_{s'}$ whenever $s \leq s'$. Assuming that $\alpha_i > 0$,*

$$\max \operatorname{supp} A_i\left(\bar{\mathbf{A}}, N, n\right) = \alpha_i \log \left(\sum_{s=N-n+1}^{N} w_s\right) + \beta_i, \text{ and}$$

$$\min \operatorname{supp} A_i\left(\bar{\mathbf{A}}, N, n\right) = \alpha_i \log \left(\sum_{s=1}^{n} w_s\right) + \beta_i.$$

Provided that condition (c) of Theorem 1.13 holds and each configuration is equally likely, *CDF $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ is as follows:*

$$G_{A_i(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{\exp\left[\frac{1}{\alpha_i}(t-\beta_i)\right] - E\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)}{\left(\operatorname{Var}\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)\right)^{1/2}}\right),$$

with $J(\cdot)$ defined in Theorem 1.13 and

$$\widehat{w}_j = \frac{\left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]_j - EW_{a,i}^{(q)}(\bar{\mathbf{A}})}{\left(N \operatorname{Var} W_{a,i}^{(q)}(\bar{\mathbf{A}})\right)^{1/2}} = \frac{\left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]_j - \frac{1}{N}}{\left(\sum_{k=1}^{N}\left(\left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]_k - \frac{1}{N}\right)^2\right)^{1/2}}.$$

Then,

$$EA_i(\bar{\mathbf{A}}, N, n) \approx \int_{-\infty}^{\infty} t\, dJ\left(\bar{\mathbf{A}}, N, n, \frac{\exp\left[\frac{1}{\alpha_i}(t-\beta_i)\right] - E\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)}{\left(\operatorname{Var}\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)\right)^{1/2}}\right)$$

and

$$\operatorname{Var} A_i(\bar{\mathbf{A}}, N, n) \approx \left[\int_{-\infty}^{\infty} t^2\, dJ\left(\bar{\mathbf{A}}, N, n, \frac{\exp\left[\frac{1}{\alpha_i}(t-\beta_i)\right] - E\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)}{\left(\operatorname{Var}\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)\right)^{1/2}}\right)\right] - \left(EA_i(\bar{\mathbf{A}}, N, n)\right)^2.$$

When agent i's action $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ nonlinearly depends on $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, provided that the action is invertible in $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, we can solve for $G_{A_i(\bar{\mathbf{A}},N,n)}(t)$ for every feasible population size, network topology, and prevalence of the attribute.

Example A.7 (Distribution of Aggregate Action, Affine Rule with Common Coefficient) *Suppose that agent i's action is: $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \alpha \widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \beta_i$, $\forall i \in \{1, \ldots, N\}$, so $a_{agg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \alpha N \widehat{f}_{avg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \mathbf{1}^T \boldsymbol{\beta}$. Then, assuming that each configuration is equally likely, $EA_{agg}(\bar{\mathbf{A}}, N, n) = \alpha n + \mathbf{1}^T \boldsymbol{\beta}$ and $\operatorname{Var} A_{agg}(\bar{\mathbf{A}}, N, n) = \alpha^2 N^2 \frac{n}{N}\left(1 - \frac{n}{N}\right) \frac{N}{N-1} N \operatorname{Var} D_w^-(\bar{\mathbf{A}})$. To compute the upper and lower bounds on the sup-*

port of $A_{agg}(\bar{\mathbf{A}}, N, n)$*, construct the ordered multiset* $\{w_s\}_{s=1}^{N}$ *from the elements of* $\mathbf{d}_w^-(\bar{\mathbf{A}})$ *so that* $w_s \leq w_{s'}$ *whenever* $s \leq s'$*. Assuming that* $\alpha > 0$*,*

$$\max \operatorname{supp} A_{agg}(\bar{\mathbf{A}}, N, n) = \alpha N \left(\sum_{s=N-n+1}^{N} w_s \right) + \mathbf{1}^T \boldsymbol{\beta}, \text{ and}$$
$$\min \operatorname{supp} A_{agg}(\bar{\mathbf{A}}, N, n) = \alpha N \left(\sum_{s=1}^{n} w_s \right) + \mathbf{1}^T \boldsymbol{\beta}.$$

Provided that condition (c) of Theorem 1.13 holds,

$$G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t) \approx J \left(\bar{\mathbf{A}}, N, n, \frac{t - \alpha N E \widehat{F}_{avg}(\bar{\mathbf{A}}, N, n) - \mathbf{1}^T \boldsymbol{\beta}}{\alpha N \left(\operatorname{Var} \widehat{F}_{avg}(\bar{\mathbf{A}}, N, n) \right)^{1/2}} \right),$$

with $J(\cdot)$ *defined in Theorem 1.13 and* $\widehat{w}_i = \frac{\left[\mathbf{d}_{\overline{w}}^-(\bar{\mathbf{A}})\right]_i - E D_{\overline{w}}^-(\bar{\mathbf{A}})}{\left(N \operatorname{Var} D_{\overline{w}}^-(\bar{\mathbf{A}})\right)^{1/2}} = \frac{\left[\mathbf{d}_{\overline{w}}^-(\bar{\mathbf{A}})\right]_i - \frac{1}{N}}{\left(\sum_{j=1}^{N} \left(\left[\mathbf{d}_{\overline{w}}^-(\bar{\mathbf{A}})\right]_j - \frac{1}{N} \right)^2 \right)^{1/2}}.$

We now relax the assumption that every agent's affine action has a common coefficient premultiplying $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$. We are able to show that we can still solve for the distribution of possible aggregate actions, $G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t)$:

Example A.8 (Distribution of Aggregate Action, Affine Rule, No Common Coefficient) *Suppose that agent i's action is:* $a_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \alpha_i \widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) + \beta_i$*,* $\forall i \in \{1, \ldots, N\}$*. The aggregate action is then:*

$$a_{agg}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \sum_{i=1}^{N} \alpha_i [\bar{\mathbf{A}}]_{i*} \mathbf{b}(N, n) + \mathbf{1}^T \boldsymbol{\beta}.$$

Define matrix $\widehat{\mathbf{A}}$ *with entry* $\left[\widehat{\mathbf{A}}\right]_{ij} = \alpha_i \left[\bar{\mathbf{A}}\right]_{ij}$ *so that* $\left[\widehat{\mathbf{A}}\right]_{i*} = \alpha_i \left[\bar{\mathbf{A}}\right]_{i*}$*. It follows that:*

$$
\begin{aligned}
a_{agg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) &= \sum_{i=1}^{N} \left[\widehat{\mathbf{A}}\right]_{i*} \mathbf{b}\left(N, n\right) + \mathbf{1}^T \boldsymbol{\beta} \\
&= \mathbf{1}^T \widehat{\mathbf{A}} \mathbf{b}\left(N, n\right) + \mathbf{1}^T \boldsymbol{\beta} \\
&= \left(\alpha_1 + \cdots + \alpha_N\right) \left(\frac{1}{\alpha_1 + \cdots + \alpha_N}\right) \mathbf{1}^T \widehat{\mathbf{A}} \mathbf{b}\left(N, n\right) + \mathbf{1}^T \boldsymbol{\beta} \\
&= \left(\alpha_1 + \cdots + \alpha_N\right) \left[\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right) + \mathbf{1}^T \boldsymbol{\beta},
\end{aligned}
$$

where $\left[\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)\right]^T = \left(\frac{1}{\alpha_1 + \cdots + \alpha_N}\right) \mathbf{1}^T \widehat{\mathbf{A}}$*. Observe that* $\left[\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)\right]^T \mathbf{1} = 1$.
Define $\widehat{f}_{avg}\left(\widehat{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right)$*, and random variable* $\widehat{F}_{avg}\left(\widehat{\mathbf{A}}, N, n\right)$ *with realization* $\widehat{f}_{avg}\left(\widehat{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right)$*. The aggregate action becomes:*

$$
a_{agg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left(\mathbf{1}^T \boldsymbol{\alpha}\right) \widehat{f}_{avg}\left(\widehat{\mathbf{A}}, \mathbf{b}, N, n\right) + \mathbf{1}^T \boldsymbol{\beta}, \text{ and}
$$

$$
A_{agg}\left(\bar{\mathbf{A}}, N, n\right) = \left(\mathbf{1}^T \boldsymbol{\alpha}\right) \widehat{F}_{avg}\left(\widehat{\mathbf{A}}, N, n\right) + \mathbf{1}^T \boldsymbol{\beta},
$$

Assuming that each configuration of the attribute is equally likely,

$$
EA_{agg}\left(\bar{\mathbf{A}}, N, n\right) = \left(\mathbf{1}^T \boldsymbol{\alpha}\right) E\widehat{F}_{avg}\left(\widehat{\mathbf{A}}, N, n\right) + \mathbf{1}^T \boldsymbol{\beta} = \left(\mathbf{1}^T \boldsymbol{\alpha}\right) \frac{n}{N} + \mathbf{1}^T \boldsymbol{\beta}, \text{ and}
$$

$$
\begin{aligned}
\operatorname{Var} A_{agg}\left(\bar{\mathbf{A}}, N, n\right) &= \left(\mathbf{1}^T \boldsymbol{\alpha}\right)^2 \operatorname{Var} \widehat{F}_{avg}\left(\widehat{\mathbf{A}}, N, n\right) \\
&= \left(\mathbf{1}^T \boldsymbol{\alpha}\right)^2 \frac{n}{N}\left(1 - \frac{n}{N}\right) \frac{N}{N-1} N \operatorname{Var} \widehat{D}_w^- \left(\widehat{\mathbf{A}}\right),
\end{aligned}
$$

where $\operatorname{Var} \widehat{D}_w^- \left(\widehat{\mathbf{A}}\right) = \frac{1}{N} \sum_{i=1}^{N} \left(\left[\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)\right]_i - \frac{1}{N}\right)^2$*. To compute the upper and lower bounds on the support of* $A_{agg}\left(\bar{\mathbf{A}}, N, n\right)$*, construct the ordered multiset* $\{w_s\}_{s=1}^N$ *from the elements of* $\widehat{\mathbf{d}}_w^- \left(\widehat{\mathbf{A}}\right)$ *so that* $w_s \leq w_{s'}$ *whenever* $s \leq s'$*. Assuming that* $\alpha_i > 0$

$\forall i \in \{1, \ldots, N\}$,

$$\max \operatorname{supp} A_{agg}\left(\bar{\mathbf{A}}, N, n\right) = \left(\mathbf{1}^T \boldsymbol{\alpha}\right) \left(\sum_{s=N-n+1}^{N} w_s\right) + \mathbf{1}^T \boldsymbol{\beta}, \text{ and}$$

$$\min \operatorname{supp} A_{agg}\left(\bar{\mathbf{A}}, N, n\right) = \left(\mathbf{1}^T \boldsymbol{\alpha}\right) \left(\sum_{s=1}^{n} w_s\right) + \mathbf{1}^T \boldsymbol{\beta}.$$

Provided that condition (c) of Theorem 1.13 holds,

$$G_{A_{agg}(\bar{\mathbf{A}},N,n)}(t) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{t - \left(\mathbf{1}^T \boldsymbol{\alpha}\right) E\widehat{F}_{avg}\left(\widehat{\mathbf{A}}, N, n\right) - \mathbf{1}^T \boldsymbol{\beta}}{\left(\mathbf{1}^T \boldsymbol{\alpha}\right)\left(\operatorname{Var} \widehat{F}_{avg}\left(\widehat{\mathbf{A}}, N, n\right)\right)^{1/2}}\right),$$

with $J(\cdot)$ *defined in Theorem 1.13 and*

$$\widehat{w}_i = \frac{\left[\widehat{\mathbf{d}}_w^-\left(\widehat{\mathbf{A}}\right)\right]_i - E\widehat{D}_w^-\left(\widehat{\mathbf{A}}\right)}{\left(N \operatorname{Var} \widehat{D}_w^-\left(\widehat{\mathbf{A}}\right)\right)^{1/2}} = \frac{\left[\widehat{\mathbf{d}}_w^-\left(\widehat{\mathbf{A}}\right)\right]_i - \frac{1}{N}}{\left(\sum_{j=1}^{N}\left(\left[\widehat{\mathbf{d}}_w^-\left(\widehat{\mathbf{A}}\right)\right]_j - \frac{1}{N}\right)^2\right)^{1/2}}.$$

Example A.9 (Distribution of Aggregate Action, Threshold Rule) *Consider a system in which the form of each agent's action is a threshold rule:*

$$a_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \begin{cases} \beta_i & \text{if } \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) \geq \alpha \\ 0 & \text{if } \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) < \alpha \end{cases}.$$

The aggregate action is as follows:

$$a_{agg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \begin{cases} \mathbf{1}^T \boldsymbol{\beta} & \text{if } \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) \geq \alpha \\ 0 & \text{if } \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) < \alpha \end{cases},$$

and

$$A_{agg}\left(\bar{\mathbf{A}}, N, n\right) = \begin{cases} \mathbf{1}^T \boldsymbol{\beta} & \text{with probability } 1 - G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(\alpha) \\ 0 & \text{with probability } G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(\alpha) \end{cases}.$$

If each configuration is equally likely,

$$EA_{agg}\left(\bar{\mathbf{A}},N,n\right)=\left(\mathbf{1}^{T}\boldsymbol{\beta}\right)\left(1-G_{\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}},N,n\right)}(\alpha)\right),\text{ and}$$

$$\begin{aligned}\operatorname{Var}A_{agg}\left(\bar{\mathbf{A}},N,n\right)&=E\left[A_{agg}\left(\bar{\mathbf{A}},N,n\right)\right]^{2}-\left(EA_{agg}\left(\bar{\mathbf{A}},N,n\right)\right)^{2}\\&=\left(\mathbf{1}^{T}\boldsymbol{\beta}\right)^{2}\left(G_{\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}},N,n\right)}(\alpha)\right)\left(1-G_{\widehat{F}^{(\infty)}\left(\bar{\mathbf{A}},N,n\right)}(\alpha)\right).\end{aligned}$$

To compute the upper and lower bounds on the support of $A_{agg}\left(\bar{\mathbf{A}},N,n\right)$*, construct the ordered multiset* $\{w_s\}_{s=1}^{N}$ *from the elements of* $\mathbf{w}_{\infty}\left(\bar{\mathbf{A}}\right)$ *so that* $w_s\leq w_{s'}$ *whenever* $s\leq s'$*. Assuming that* $\mathbf{1}^{T}\boldsymbol{\beta}>0$*,*

$$\max\operatorname{supp}A_{agg}\left(\bar{\mathbf{A}},N,n\right)=\begin{cases}\mathbf{1}^{T}\boldsymbol{\beta} & \text{if } \sum_{s=N-n+1}^{N}w_s\geq\alpha\\ 0 & \text{otherwise}\end{cases},\text{ and}$$

$$\min\operatorname{supp}A_{agg}\left(\bar{\mathbf{A}},N,n\right)=\begin{cases}0 & \text{if } \sum_{s=1}^{n}w_s<\alpha\\ \mathbf{1}^{T}\boldsymbol{\beta} & \text{otherwise}\end{cases}.$$

Provided that condition (c) of Theorem 1.13 holds,

$$A_{agg}\left(\bar{\mathbf{A}},N,n\right)\approx\begin{cases}\mathbf{1}^{T}\boldsymbol{\beta} & \text{with probability } 1-J\left(\bar{\mathbf{A}},N,n,\frac{\alpha-E\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}{\left(\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)\right)^{1/2}}\right)\\ 0 & \text{with probability } J\left(\bar{\mathbf{A}},N,n,\frac{\alpha-E\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}{\left(\operatorname{Var}\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)\right)^{1/2}}\right)\end{cases},$$

with $J(\cdot)$ *defined in Theorem 1.13 and* $\widehat{w}_i=\frac{[\mathbf{w}_{\infty}(\bar{\mathbf{A}})]_i-EW_{\infty}(\bar{\mathbf{A}})}{\left(N\operatorname{Var}W_{\infty}(\bar{\mathbf{A}})\right)^{1/2}}=\frac{[\mathbf{w}_{\infty}(\bar{\mathbf{A}})]_i-\frac{1}{N}}{\left(\sum_{j=1}^{N}\left([\mathbf{w}_{\infty}(\bar{\mathbf{A}})]_j-\frac{1}{N}\right)^2\right)^{1/2}}$.

Example A.10 (First Moment of Aggregate Action, Threshold Rule) *Consider a system in which agent i's action,* $\forall i\in\{1,\ldots,N\}$*, takes the following form:*

$$a_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)=\begin{cases}\beta_i & \text{if } \widehat{f}_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)\geq\alpha\\ 0 & \text{if } \widehat{f}_i\left(\bar{\mathbf{A}},\mathbf{b},N,n\right)<\alpha\end{cases}.$$

Random variable $A_i\left(\bar{\mathbf{A}}, N, n\right)$ *takes the following values:*

$$A_i\left(\bar{\mathbf{A}}, N, n\right) = \begin{cases} \beta_i & \textit{with probability } 1 - G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(\alpha) \\ 0 & \textit{with probability } G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(\alpha) \end{cases}.$$

As a result,

$$EA_{agg}\left(\bar{\mathbf{A}}, N, n\right) = \sum_{i=1}^{N} EA_i\left(\bar{\mathbf{A}}, N, n\right) = \sum_{i=1}^{N} \beta_i \left(1 - G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(\alpha)\right).$$

Provided that condition (c) from Theorem 1.13 holds and each configuration of the attribute is equally likely to occur,

$$G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(\alpha) \approx J\left(\bar{\mathbf{A}}, N, n, \frac{\alpha - E\widehat{F}_i\left(\bar{\mathbf{A}}, N, n\right)}{\left(\mathrm{Var}\,\widehat{F}_i\left(\bar{\mathbf{A}}, N, n\right)\right)^{1/2}}\right),$$

with $J(\cdot)$ *defined in Theorem 1.13 and* $\widehat{w}_j = \frac{\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_j - EW_{a,i}(\bar{\mathbf{A}})}{\left(N\,\mathrm{Var}\,W_{a,i}(\bar{\mathbf{A}})\right)^{1/2}} = \frac{\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_j - \frac{1}{N}}{\left(\sum_{k=1}^{N}\left(\left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]_k - \frac{1}{N}\right)^2\right)^{1/2}}$*,* *so*

$$EA_{agg}\left(\bar{\mathbf{A}}, N, n\right) \approx \sum_{i=1}^{N} \beta_i \left[1 - J\left(\bar{\mathbf{A}}, N, n, \frac{\alpha - E\widehat{F}_i\left(\bar{\mathbf{A}}, N, n\right)}{\left(\mathrm{Var}\,\widehat{F}_i\left(\bar{\mathbf{A}}, N, n\right)\right)^{1/2}}\right)\right].$$

We have thus solved for the first moment of the aggregate action when each agent follows a threshold rule.

A.11 Proofs

Proof of Theorem 1.1

Since $\bar{\mathbf{A}}$ is an $N \times N$ nonnegative primitive matrix, by the Perron-Frobenius Theorem, there exists (1) a positive real eigenvalue $\lambda_1 > |\lambda_i|$ for all other $\lambda_i \neq \lambda_1$, (2) an associated left eigenvector $\mathbf{w}^T \in \mathbb{R}_{++}^{1\times N}$, and (3) an associated right eigenvector $\mathbf{v} \in \mathbb{R}_{++}^{N\times 1}$ both unique up to a constant. Consider the following lemma:

Lemma A.1 (Seneta (1981), Corollary 1 to Theorem 1.1) *For a primitive matrix* $\mathbf{X}$*,*

$$\min_{i\in\{1,\ldots,N\}} \sum_{j=1}^{N} [\mathbf{X}]_{ij} \leq \lambda_1 \leq \max_{i\in\{1,\ldots,N\}} \sum_{j=1}^{N} [\mathbf{X}]_{ij} .$$

Since $\bar{\mathbf{A}}$ is row-stochastic, $\lambda_1 = 1$ by Lemma A.1.

Now let the distinct eigenvalues of $\bar{\mathbf{A}}$ be $\lambda_1, \lambda_2, \ldots, \lambda_r$, $r \leq N$, with $\lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_r|$ and $|\cdot|$ denoting the eigenvalue modulus. If two eigenvalues share the same modulus, the one with the weakly greater algebraic multiplicity receives the lower integer index i. The next lemma characterizes the convergence behavior of $\bar{\mathbf{A}}^q$:

Lemma A.2 (Seneta (1981), Theorem 1.2) *For a primitive matrix* $\mathbf{X}$*, if* $\lambda_2 \neq 0$*, then as* $q \to \infty$*,*

$$\mathbf{X}^q = \lambda_1^q \mathbf{v}\mathbf{w}^T + \mathcal{O}\left(q^{m_2-1} |\lambda_2|^q\right)$$

element-wise, where m_2 *is the algebraic multiplicity of* λ_2*; otherwise, if* $\lambda_2 = 0$*, for* $q \geq N-1$, $\mathbf{X}^q = \lambda_1^q \mathbf{v}\mathbf{w}^T$.

By Lemma A.2, $\lim_{q\to\infty} \bar{\mathbf{A}}^q = \mathbf{v}\mathbf{w}^T$.

From the row-stochasticity of $\bar{\mathbf{A}}$, $\mathbf{v} = \alpha\mathbf{1}$ for some non-zero constant α, so $\lim_{q\to\infty} \bar{\mathbf{A}}^q = \alpha\mathbf{1}\mathbf{w}^T$. Since row-stochasticity is preserved under matrix multiplication, $\lim_{q\to\infty} \bar{\mathbf{A}}^q$ must also be row-stochastic, so

$$\lim_{q\to\infty} \bar{\mathbf{A}}^q = \begin{pmatrix} \text{---} & \alpha\mathbf{w}^T & \text{---} \\ & \vdots & \\ \text{---} & \alpha\mathbf{w}^T & \text{---} \end{pmatrix}$$

with $\alpha \mathbf{w}^T \mathbf{1} = 1$. Set $\mathbf{w}_\infty^T = \alpha \mathbf{w}^T$. Then

$$\lim_{q \to \infty} \bar{\mathbf{A}}^q = \begin{pmatrix} \text{——} & \mathbf{w}_\infty^T & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{w}_\infty^T & \text{——} \end{pmatrix}.$$

Therefore, $\lim_{q \to \infty} w_{ij}^{(q)} = \left[\mathbf{w}_\infty^T\right]_j$ exists, $\mathbf{w}_\infty^T \mathbf{1} = 1$, and $\mathbf{w}_\infty^T \bar{\mathbf{A}} = \mathbf{w}_\infty^T$, with $\left(\mathbf{w}_\infty^T, 1\right)$ the unique dominant left eigenpair given the constraint $\mathbf{w}_\infty^T \mathbf{1} = 1$. $\square$

Proof of Theorem 1.2

Since graph $\mathcal{G}\left(\bar{\mathbf{A}}\right)$ is undirected and all non-zero elements within every row of $\bar{\mathbf{A}}$ have the same value, either $[\bar{\mathbf{A}}]_{ij} = \frac{1}{d_i}$ or $[\bar{\mathbf{A}}]_{ij} = 0$. Conjecture the solution $\mathbf{w}_\infty = \frac{\mathbf{d}}{\mathbf{1}^T \mathbf{d}}$ to $\mathbf{w}_\infty^T = \mathbf{w}_\infty^T \bar{\mathbf{A}}$, where $\mathbf{d}$ is the degree vector for graph $\mathcal{G}\left(\bar{\mathbf{A}}\right)$. Then,

$$\mathbf{w}_\infty^T \bar{\mathbf{A}} = \left[\frac{\mathbf{d}}{\mathbf{1}^T \mathbf{d}}\right]^T \bar{\mathbf{A}} = \left(\text{——} \quad \sum_{i=1}^N \frac{d_i [\bar{\mathbf{A}}]_{ij}}{\mathbf{1}^T \mathbf{d}} \quad \text{——}\right) = \left(\text{——} \quad \frac{d_j}{\mathbf{1}^T \mathbf{d}} \quad \text{——}\right) = \left[\frac{\mathbf{d}}{\mathbf{1}^T \mathbf{d}}\right]^T$$
$$= \mathbf{w}_\infty^T.$$

We next demonstrate that $\mathbf{1}^T \mathbf{w}_\infty$:

$$\mathbf{1}^T \mathbf{w}_\infty = \mathbf{1}^T \left[\frac{\mathbf{d}}{\mathbf{1}^T \mathbf{d}}\right] = 1. \quad \square$$

Proof of Theorem 1.3

Since graph $\mathcal{G}\left(\mathbf{A}\right)$ is directed and all non-zero elements within every row of $\bar{\mathbf{A}}$ have the same value, either $[\bar{\mathbf{A}}]_{ij} = \frac{1}{d_i^+}$ or $[\bar{\mathbf{A}}]_{ij} = 0$. Conjecture the solution $\mathbf{w}_\infty = \frac{\mathbf{d}^+}{\mathbf{1}^T \mathbf{d}^+}$

to $\mathbf{w}_\infty^T = \mathbf{w}_\infty^T \bar{\mathbf{A}}$, where $\mathbf{d}^+$ is the vector of out-degrees for graph $\mathcal{G}(\mathbf{A})$. Then,

$$\begin{aligned}\mathbf{w}_\infty^T \bar{\mathbf{A}} = \left[\frac{\mathbf{d}^+}{\mathbf{1}^T\mathbf{d}^+}\right]^T \bar{\mathbf{A}} = \left(\text{——} \quad \sum_{i=1}^N \frac{d_i^+ [\bar{\mathbf{A}}]_{ij}}{\mathbf{1}^T\mathbf{d}^+} \quad \text{——}\right) &= \left(\text{——} \quad \frac{d_j^-}{\mathbf{1}^T\mathbf{d}^+} \quad \text{——}\right) \\ &= \left(\text{——} \quad \frac{d_j^+}{\mathbf{1}^T\mathbf{d}^+} \quad \text{——}\right) \\ &= \left[\frac{\mathbf{d}^+}{\mathbf{1}^T\mathbf{d}^+}\right]^T = \mathbf{w}_\infty^T,\end{aligned}$$

with $d_i^- = d_i^+$ because graph $\mathcal{G}$ is Eulerian. $\mathbf{1}^T\mathbf{w}_\infty = \mathbf{1}^T \left[\frac{\mathbf{d}^+}{\mathbf{1}^T\mathbf{d}^+}\right] = 1$, so $\mathbf{1}^T\mathbf{w}_\infty = 1$.

□

Proof of Theorem 1.4

By viewing $\mathbf{w}_\infty$ as mathematically analogous to the stationary distribution of a simple random walk on a random digraph, the statements in this Theorem follow from Theorem 2, Lemma 14, and the Remarks of Theorem 2 from Cooper and Frieze (2012). When $\chi(N) = 1 + \kappa$, $\kappa > 0$, or $(\chi(N) - 1)\log N = \omega(\log\log N)$, w.h.p. $\iota_i = o\left(d_i^-\right)$, so it follows that w.h.p. $\mathbf{w}_\infty \sim \frac{\mathbf{d}^-}{E[|\mathcal{E}|]}$. Similarly, by Lemma 14 of Cooper and Frieze (2012), w.h.p. $\iota_i = o\left(d_i^-\right)$ for $N - o\left(N^{1/4}\right)$ nodes, so w.h.p. $\mathbf{w}_\infty \sim \frac{\mathbf{d}^-}{E[|\mathcal{E}|]}$ for $N - o\left(N^{1/4}\right)$ nodes. □

Proof of Theorem 1.5

The scalar $x = \mathbf{w}^T\mathbf{b}$ is invariant to configuration if and only if $\mathbf{w}^T\mathbf{b}(N, n) = \mathbf{w}^T\mathbf{b}'(N, n)$ for all $\mathbf{b}(N, n), \mathbf{b}'(N, n) \in \mathcal{B}(N, n)$, with this relation holding for each integer $n \in [0, N]$. Let $n = 1$, and define $\mathbf{e}_i$ to be the i^{th} unit vector whose i^{th} element equals 1 and all other elements equal zero. Then $\mathbf{w}^T\mathbf{b}(N, 1) = \mathbf{w}^T\mathbf{b}'(N, 1)$ if and only if $\mathbf{w}^T\mathbf{e}_i = \mathbf{w}^T\mathbf{e}_j$, that is, if and only if $w_i = w_j$ for all $i, j \in \{1, \ldots, N\}$. Since $\mathbf{1}^T\mathbf{w} = 1$, $w_i = \frac{1}{N}$ for all $i \in \{1, \ldots, N\}$. Given that $\mathbf{w} = \frac{1}{N}\mathbf{1}$ when x is invariant to configuration, $x = \mathbf{w}^T\mathbf{b} = \frac{1}{N}\mathbf{1}^T\mathbf{b} = \frac{n}{N}$. □

Proof of Corollary 1.1

This corollary follows immediately from Theorem 1.5. For part (1), using the objects defined in Theorem 1.5, first replace x with $\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ (respectively $\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$) and next replace $\mathbf{w}$ with $\mathbf{w}_{a,i}$ (resp. $\mathbf{w}_{a,i}^{(q)}$). For part (2), $\widehat{\mathbf{f}}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ (resp. $\widehat{\mathbf{f}}^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$) is invariant to configuration if and only if $\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ (resp. $\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$) is invariant to configuration for all $i \in \{1, \ldots, N\}$, that is, if and only if $[\mathbf{w}_{a,i}]_j = \frac{1}{N}$ (resp. $\left[\mathbf{w}_{a,i}^{(q)}\right]_j = \frac{1}{N}$) for all $i, j \in \{1, \ldots, N\}$. For part (3), first replace x with $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ (resp. $\widehat{f}_{avg}^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$) and next replace $\mathbf{w}$ with $\mathbf{d}_w^-$ (resp. $\mathbf{d}_w^{-\,(q)}$). For part (4), first replace x with $\widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ and next replace $\mathbf{w}$ with $\mathbf{w}_\infty$. □

Proof of Theorem 1.6

This proof employs the statements of Corollary 1.1. For part (1), $\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $\mathbf{w}_{a,i} = \frac{1}{N}\mathbf{1}$. Since $\mathbf{w}_{a,i}^T \equiv [\bar{\mathbf{A}}]_{i*}$, $\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $[\bar{\mathbf{A}}]_{i*} = \frac{1}{N}\mathbf{1}^T$. Meanwhile, $\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $\mathbf{w}_{a,i}^{(q)} = \frac{1}{N}\mathbf{1}$. Since $\left[\mathbf{w}_{a,i}^{(q)}\right]^T \equiv [\bar{\mathbf{A}}^q]_{i*}$, $\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $[\bar{\mathbf{A}}^q]_{i*} = \frac{1}{N}\mathbf{1}^T$. For part (2), $\widehat{\mathbf{f}}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $[\bar{\mathbf{A}}]_{i*} = \frac{1}{N}\mathbf{1}^T$ for all $i \in \{1, \ldots, N\}$, that is, if and only if $\bar{\mathbf{A}} = \frac{1}{N}\mathbf{1}\mathbf{1}^T$. Meanwhile, $\widehat{\mathbf{f}}^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $[\bar{\mathbf{A}}^q]_{i*} = \frac{1}{N}\mathbf{1}^T$ for all $i \in \{1, \ldots, N\}$, that is, if and only if $\bar{\mathbf{A}}^q = \frac{1}{N}\mathbf{1}\mathbf{1}^T$. For part (3), $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $\mathbf{d}_w^- = \frac{1}{N}\mathbf{1}$. Since $\mathbf{d}_w^- \equiv \frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1}$, $\widehat{f}_{avg}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)$ is invariant to configuration if and only if $\frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1} = \frac{1}{N}\mathbf{1}$, that is, if and only if $\bar{\mathbf{A}}$

is doubly stochastic. Meanwhile, $\widehat{f}_{avg}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is invariant to configuration if and only if $\mathbf{d}_w^{-(q)} = \frac{1}{N}\mathbf{1}$. Since $\mathbf{d}_w^{-(q)} \equiv \frac{1}{N}[\bar{\mathbf{A}}^q]^T \mathbf{1}$, $\widehat{f}_{avg}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is invariant to configuration if and only if $\frac{1}{N}[\bar{\mathbf{A}}^q]^T \mathbf{1} = \frac{1}{N}\mathbf{1}$, that is, if and only if $\bar{\mathbf{A}}^q$ is doubly stochastic. For part (4), $\widehat{f}^{(\infty)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is invariant to configuration if and only if $\mathbf{w}_\infty = \frac{1}{N}\mathbf{1}$. Since $\mathbf{w}_\infty^T$ is the left eigenvector associated with the unique dominant unit eigenvalue of $\bar{\mathbf{A}}$, so that $\mathbf{w}_\infty^T = \mathbf{w}_\infty^T\bar{\mathbf{A}}$, the vector $\mathbf{w}_\infty = \frac{1}{N}\mathbf{1}$ if and only if $\mathbf{1}^T = \mathbf{1}^T\bar{\mathbf{A}}$, that is, if and only if $\bar{\mathbf{A}}$ is doubly stochastic. □

Proof of Theorem 1.7

Construct the set of $N+1$ vectors $\{\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_n, \ldots, \mathbf{b}_N\}$, for which $[\mathbf{b}_n]_i = 1$ for $i \leq n$ and $[\mathbf{b}_n]_i = 0$ for $n+1 \leq i \leq N$. For each integer $n \in [0, N]$, also define the set $\mathcal{S}_n$ of all $N \times N$ permutation matrices $\mathbf{S}_{nx}, \mathbf{S}_{ny} \in \mathcal{S}_n$ for which $\mathbf{S}_{nx}\mathbf{b}_n \neq \mathbf{b}_n$ and $\mathbf{S}_{nx}\mathbf{b}_n \neq \mathbf{S}_{ny}\mathbf{b}_n$.

Lemma A.3 *Unordered multiset* $\left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^N$ *(respectively* $\left\{\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^N$*) is invariant to configuration if and only if, for each* $\mathbf{S}_{nx} \in \mathcal{S}_n$ *and for every* $n \in [0, N]$*, there corresponds some permutation matrix* $\mathbf{R}$ *such that* $\bar{\mathbf{A}}\mathbf{b}_n = \mathbf{R}\bar{\mathbf{A}}\mathbf{S}_{nx}\mathbf{b}_n$ *(respectively such that* $\bar{\mathbf{A}}^q\mathbf{b}_n = \mathbf{R}\bar{\mathbf{A}}^q\mathbf{S}_{nx}\mathbf{b}_n$*).*

Proof. Unordered multiset $\left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^N$ is invariant to configuration if and only if, for every configuration $\mathbf{b}' \in \mathcal{B}(N, n)$, there exists an $N \times N$ permutation matrix $\mathbf{R}$ such that $\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \mathbf{R}\widehat{\mathbf{f}}(\bar{\mathbf{A}}, \mathbf{b}', N, n)$, or equivalently, $\bar{\mathbf{A}}\mathbf{b} = \mathbf{R}\bar{\mathbf{A}}\mathbf{b}'$. Now, generate each configuration in the set $\mathcal{B}(N, n)$ by introducing the vector $\mathbf{b}_n \in \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ and defining the set $\mathcal{S}_n$ of all $N \times N$ permutation matrices $\mathbf{S}_{nx}, \mathbf{S}_{ny} \in \mathcal{S}_n$, for which $\mathbf{S}_{nx}\mathbf{b}_n \neq \mathbf{b}_n$ and $\mathbf{S}_{nx}\mathbf{b}_n \neq \mathbf{S}_{ny}\mathbf{b}_n$. Then, for every $\mathbf{b}' \in \mathcal{B}(N, n)$, there exists some $\mathbf{S}_{nx} \in \mathcal{S}_n$ for which $\mathbf{b}' = \mathbf{S}_{nx}\mathbf{b}_n$. Multiset $\left\{\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)\right\}_{i=1}^N$ is thus invariant to configuration if and only if, for each $\mathbf{S}_{nx} \in \mathcal{S}_n$

and for every $n \in [0, N]$, there corresponds some permutation matrix $\mathbf{R}$ such that $\bar{\mathbf{A}}\mathbf{b}_n = \mathbf{R}\bar{\mathbf{A}}\mathbf{S}_{nx}\,\mathbf{b}_n$. Replace matrix $\bar{\mathbf{A}}$ with $\bar{\mathbf{A}}^q$ to obtain the result that multiset $\left\{\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$ is invariant to configuration if and only if, for each $\mathbf{S}_{nx} \in \mathcal{S}_n$ and every $n \in [0, N]$, there corresponds some permutation matrix $\mathbf{R}$ such that $\bar{\mathbf{A}}^q\mathbf{b}_n = \mathbf{R}\bar{\mathbf{A}}^q\mathbf{S}_{nx}\,\mathbf{b}_n$. ■

From Lemma A.3, multiset $\left\{\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$ is invariant to configuration if and only if there exists a permutation matrix $\mathbf{R}$ such that

$$\bar{\mathbf{A}}\mathbf{b}_n = \mathbf{R}\bar{\mathbf{A}}\mathbf{S}_{nx}\mathbf{b}_n \tag{A.1}$$

for each $\mathbf{S}_{nx} \in \mathcal{S}_n$ and for every $n \in [0, N]$. On the left-hand side of Equation A.1, $\bar{\mathbf{A}}\mathbf{b}_n$ is the row sum of the first n columns of $\bar{\mathbf{A}}$. On the right-hand side of Equation A.1, $\mathbf{S}_{nx}$ permutes the columns of $\bar{\mathbf{A}}$, so $\bar{\mathbf{A}}\mathbf{S}_{nx}\mathbf{b}_n$ is the row sum of a different set of n column vectors of $\bar{\mathbf{A}}$. With matrix $\mathbf{R}$ permuting the rows of $\bar{\mathbf{A}}\mathbf{S}_{nx}$, $\bar{\mathbf{A}}\mathbf{b}_n = \mathbf{R}\bar{\mathbf{A}}\mathbf{S}_{nx}\mathbf{b}_n$ when the multiset of elements formed from the row sum of the first n columns of $\bar{\mathbf{A}}$ equals the multiset of elements formed from the row sum of a different group of n columns of $\bar{\mathbf{A}}$. Now considering all allowable permutation matrices $\mathbf{S}_{nx}$, multiset $\left\{\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$ is invariant to configuration if and only if the row sum of any n columns of $\bar{\mathbf{A}}$ has the same multiset of elements, and this property holds for every integer $n \in [1, N]$.

It turns out that it is redundant to specify that this condition must hold for every $n \in [1, N] \subseteq \mathbb{Z}_+$. Since $\bar{\mathbf{A}}$ is row-stochastic, the row sum of any n column vectors of $\bar{\mathbf{A}}$, $n \in [1, N-1]$, has the same multiset of elements whenever the row sum of any $N - n$ column vectors of $\bar{\mathbf{A}}$ has the same multiset of elements. Furthermore, the row sum of N column vectors of $\bar{\mathbf{A}}$ is always equal to $\mathbf{1}$. Therefore, we eliminate such redundancies in the required set of integers n for which this condition must hold and state that multiset $\left\{\widehat{f}_i\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right)\right\}_{i=1}^N$ is invariant to configuration if and only if the row sum of any n columns of $\bar{\mathbf{A}}$ has the same multiset of elements for every integer $n \in \left[1, \lfloor \frac{N}{2} \rfloor\right]$. Next, since $\bar{\mathbf{A}}$ is row-stochastic, if the row sum of any n column vectors of $\bar{\mathbf{A}}$ has the same multiset of elements for every integer $n \in \left[1, \lfloor \frac{N}{2} \rfloor\right]$, $\bar{\mathbf{A}}$ must be doubly stochastic. For $n = 1$, every column of $\bar{\mathbf{A}}$ must have the same multiset of elements, so every column of $\bar{\mathbf{A}}$ must have the same sum α. The sum of all matrix elements in $\bar{\mathbf{A}}$ is N, by the row-stochasticity of $\bar{\mathbf{A}}$, so $N\alpha = N$, that is, $\alpha = 1$. $\square$

Proof of Theorem 1.8

First method of proof: $EX\left(\bar{\mathbf{A}}, N, n\right) = E\left[\mathbf{w}^T \mathbf{B}\left(N, n\right)\right]$, where $\mathbf{B}\left(N, n\right)$ is a random vector whose elements are $B_i \sim \text{Bern}\left(\frac{n}{N}\right), i \in \{1, \ldots, N\}$. Therefore,

$$EX\left(\bar{\mathbf{A}}, N, n\right) = E\left(w_1 B_1 + w_2 B_2 + \ldots + w_N B_N\right) = \sum_{i=1}^{N} w_i E B_i = \frac{n}{N}. \quad \square$$

Second method of proof:

$$\begin{aligned}
EX\left(\bar{\mathbf{A}}, N, n\right) &= \frac{1}{\left|\mathcal{B}\left(N, n\right)\right|} \sum_{\mathbf{b}(N,n) \in \mathcal{B}(N,n)} \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right) \\
&= \frac{1}{\left|\mathcal{B}\left(N, n\right)\right|} \sum_{i=1}^{N} \left|\left\{\mathbf{b} \in \mathcal{B}\left(N, n\right) : b_i = 1\right\}\right| w_i \\
&= \frac{\left|\left\{\mathbf{b} \in \mathcal{B}\left(N, n\right) : b_i = 1\right\}\right|}{\left|\mathcal{B}\left(N, n\right)\right|} \sum_{i=1}^{N} w_i = \frac{\binom{N}{n} \times \frac{n}{N}}{\binom{N}{n}} = \frac{n}{N}.
\end{aligned}$$

By symmetry, $\left|\left\{\mathbf{b} \in \mathcal{B}\left(N, n\right) : b_i = 1\right\}\right| = \left|\left\{\mathbf{b} \in \mathcal{B}\left(N, n\right) : b_j = 1\right\}\right|$, $\forall i, j \in \{1, \ldots, N\}$. $\square$

Proof of Corollary 1.2

Replace the pair $\left(X, \mathbf{w}\right)$ in the proof of Theorem 1.8 with each random variable specified in the statement of Corollary 1.2 and its associated vector of weights. For example, to demonstrate that $E\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}$, replace $(X, \mathbf{w})$ in the proof of Theorem 1.8 with $\left(\widehat{F}_{avg}\left(\bar{\mathbf{A}}, N, n\right), \mathbf{d}_{w}^{-}\left(\bar{\mathbf{A}}\right)\right)$. Corollary 1.2 then follows. □

Proof of Theorem 1.9

First demonstrating that $\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W\left(\bar{\mathbf{A}}\right)\right)$:

$\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n\right) = \operatorname{Var}\left[\mathbf{w}^{T}\mathbf{B}\left(N, n\right)\right]$, where $\mathbf{B}\left(N, n\right)$ is a random vector whose elements are $B_i \sim \operatorname{Bern}\left(\frac{n}{N}\right)$, for all $i \in \{1, \ldots, N\}$. Therefore,

$$\begin{aligned}
\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n\right) &= \operatorname{Var}\left(w_1 B_1 + w_2 B_2 + \cdots + w_N B_N\right) \\
&= \sum_{i=1}^{N} \operatorname{Var}\left(w_i B_i\right) + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \operatorname{Cov}\left(w_i B_i, w_j B_j\right) \\
&= \sum_{i=1}^{N} w_i^2 \operatorname{Var} B_i + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_i w_j \operatorname{Cov}\left(B_i, B_j\right) \\
&= \left(\operatorname{Var} B_i\right) \sum_{i=1}^{N} w_i^2 + \left(E\left[B_i B_j\right] - \left(E B_i\right)\left(E B_j\right)\right) \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_i w_j \\
&= \frac{n}{N}\left(1 - \frac{n}{N}\right) \sum_{i=1}^{N} w_i^2 + \left[\left(\frac{n}{N}\right)\left(\frac{n-1}{N-1}\right) - \left(\frac{n}{N}\right)^2\right] \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_i w_j \\
&= \frac{n}{N}\left(1 - \frac{n}{N}\right) \sum_{i=1}^{N} w_i^2 - \frac{n\left(1 - \frac{n}{N}\right)}{N\left(N-1\right)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_i w_j \\
&= \frac{n}{N}\left(1 - \frac{n}{N}\right) \times \left(\sum_{i=1}^{N} w_i^2 - \frac{1}{N-1} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_i w_j\right).
\end{aligned}$$

$$\text{Now, } \operatorname{Var} W = \frac{1}{N}\sum_{i=1}^{N}\left(w_i - \frac{\sum_{j=1}^{N} w_j}{N}\right)^2 = \frac{1}{N}\left(\sum_{i=1}^{N} w_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} w_i\right)^2\right)$$

$$= \frac{1}{N}\left(\sum_{i=1}^{N} w_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} w_i^2 + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} w_i w_j\right)\right)$$

$$= \frac{1}{N}\frac{N-1}{N}\left[\sum_{i=1}^{N} w_i^2 - \frac{1}{N-1}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} w_i w_j\right],$$

$$\text{so } \operatorname{Var} X\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N \operatorname{Var} W\right).$$

Next demonstrating that $\operatorname{Var} X\left(\bar{\mathbf{A}}, N, n\right) \to 0$ at rate N^{-1} as $N \to \infty$ assuming $\operatorname{Var} W < \infty$:

Consider the graph $\mathcal{G}\left(\bar{\mathbf{A}}_1\right) = \left(\mathcal{V}\left(\bar{\mathbf{A}}_1\right), \mathcal{E}\left(\bar{\mathbf{A}}_1\right)\right)$, $\left|\mathcal{V}\left(\bar{\mathbf{A}}_1\right)\right| = N_1$, corresponding to the row-stochastic matrix $\bar{\mathbf{A}}_1$ with vector $\mathbf{w}_1$. Construct replica graphs $\mathcal{G}\left(\bar{\mathbf{A}}_2\right), \ldots, \mathcal{G}\left(\bar{\mathbf{A}}_K\right)$, with $\bar{\mathbf{A}}_1 = \bar{\mathbf{A}}_2 = \cdots = \bar{\mathbf{A}}_K$ and $\mathbf{w}_1 = \mathbf{w}_2 = \cdots = \mathbf{w}_K$. Next define $\mathcal{G}\left(\bar{\mathbf{A}}\right) = \left(\mathcal{V}\left(\bar{\mathbf{A}}\right), \mathcal{E}\left(\bar{\mathbf{A}}\right)\right)$, $\left|\mathcal{V}\left(\bar{\mathbf{A}}\right)\right| = N_1 K \equiv N$, where $\bar{\mathbf{A}} = \operatorname{diag}\left(\bar{\mathbf{A}}_1, \bar{\mathbf{A}}_2, \ldots, \bar{\mathbf{A}}_K\right)$ is an $N \times N = N_1 K \times N_1 K$ block diagonal matrix, and $\mathbf{w} = \left(\mathbf{w}_1 \mathbf{w}_2 \cdots \mathbf{w}_K\right)^T / K$ is an $N \times 1 = N_1 K \times 1$ vector of weights, with individual elements normalized by K to ensure that $\mathbf{1}^T \mathbf{w} = 1$. Then, letting $B_{ij} \sim \operatorname{Bern}\left(\frac{n}{N_1}\right)$ for $i \in \{1, \ldots, N_1\}$ and $j \in \{1, \ldots, K\}$, with B_{ij}, B_{gr} independent for all $j \neq r$,

$$\operatorname{Var} X\left(\bar{\mathbf{A}}, N_1 K, nK\right) = \operatorname{Var}\left(\frac{1}{K}\left(w_{11}B_{11} + \cdots + w_{N_1 1}B_{N_1 1} + w_{12}B_{12} + \cdots + w_{N_1 2}B_{N_1 2}\right.\right.$$
$$\left.\left. + \cdots + w_{1K}B_{1K} + \cdots + w_{N_1 K}B_{N_1 K}\right)\right)$$
$$= \frac{1}{K^2} K \operatorname{Var}\left(w_{11}B_1 + \cdots + w_{1N}B_N\right) = \frac{1}{K}\operatorname{Var} X\left(\bar{\mathbf{A}}_1, N_1, n\right) = \frac{N_1}{N}\operatorname{Var} X\left(\bar{\mathbf{A}}_1, N_1, n\right),$$

so $\operatorname{Var} X\left(\bar{\mathbf{A}}, N_1 K, nK\right) \to 0$ at rate N^{-1} as $N \to \infty$. □

Proof of Theorem 1.10

$x\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right) = \sum_{i\in\{1,\ldots,N\} \text{ s.t. } [\mathbf{b}]_i=1} [\mathbf{w}]_i$. Therefore,

$$\min \text{supp } X\left(\bar{\mathbf{A}}, N, n\right) = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}_*\left(N, n\right) = \sum_{i=1}^{n} w_s,$$

where w_s is the s^{th} smallest element in $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ in the ordered multiset $\{w_s\}_{s=1}^N$ and $\mathbf{b}_*\left(N, n\right)$ is defined so that $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}_*\left(N, n\right) \leq \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right)$ for all $\mathbf{b}\left(N, n\right) \in \mathcal{B}\left(N, n\right)$. Meanwhile,

$$\max \text{supp } X\left(\bar{\mathbf{A}}, N, n\right) = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}^*\left(N, n\right) = \sum_{i=N-n+1}^{N} w_s,$$

where w_s is the s^{th} smallest element of $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ listed in the ordered multiset $\{w_s\}_{s=1}^N$ and $\mathbf{b}^*\left(N, n\right)$ is defined so that $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}^*\left(N, n\right) \geq \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T \mathbf{b}\left(N, n\right)$ for all $\mathbf{b}\left(N, n\right) \in \mathcal{B}\left(N, n\right)$. $\square$

Proof of Theorem 1.11

The proof of Theorem 1.11 makes use of the following result from Erdös and Rényi (1959), with notation modified for the present work:

Lemma A.4 (Erdös and Rényi (1959), Theorem 1) *Consider the infinite triangular matrix of real elements*

$$\begin{matrix} w'_{11} & & & & & & \\ w'_{21} & w'_{22} & & & & & \\ . & . & . & & & & \\ . & . & . & . & & & \\ . & . & . & . & . & & \\ w'_{N1} & w'_{N2} & . & . & . & w'_{NN} & \\ . & . & . & . & . & . & . \end{matrix}$$

with $\mathbf{w}_N^{'}$ *denoting the* N^{th} *row of the matrix and* $\sum_{j=1}^{N} w_{Nj}^{'} = 0$. *For any real value* t, *determine* $T\left(\mathbf{w}_N^{'}, N, n, t\right)$, *that is, the total number of sums*

$$y\left(\mathbf{w}_N^{'}, N, n\right) = w_{Ni_1}^{'} + w_{Ni_2}^{'} + \cdots + w_{Ni_n}^{'}, \qquad 1 \leq i_1 < i_2 < \cdots < i_n \leq N,$$

whose value does not exceed $t\sigma\left(\mathbf{w}_N^{'}, N, n\right) \equiv t\sqrt{\frac{n}{N}\left(1-\frac{n}{N}\right)\sum_{j=1}^{N} w_{Nj}^{'2}}$. *Let CDF* $G_{\frac{Y\left(\mathbf{w}_N^{'},N,n\right)}{\sigma\left(\mathbf{w}_N^{'},N,n\right)}}(t) = \frac{T\left(\mathbf{w}_N^{'},N,n,t\right)}{\binom{N}{n}}$. *With*

$$\kappa\left(\mathbf{w}_N^{'}, N, n, \epsilon\right) \equiv \frac{1}{\sum_{j=1}^{N} w_{Nj}^{'2}} \sum_{\substack{j\in\{1,\ldots,N\} \text{ s.t.} \\ \left|w_{Nj}^{'}\right| > \epsilon\sigma\left(\mathbf{w}_N^{'},N,n\right)}} w_{Nj}^{'2}$$

if $\lim_{N\to\infty} \kappa\left(\mathbf{w}_N^{'}, N, n, \epsilon\right) = 0$ *for any* $\epsilon > 0$, *then* $\lim_{N\to\infty} G_{\frac{Y}{\sigma}}(t) = \Phi(t)$ *for any real* t, *where* $\Phi(\cdot)$ *denotes the standard normal CDF.*

For a given population size N, set $\mathbf{w}_N^{'} = \mathbf{w}_N(\bar{\mathbf{A}}) - \frac{1}{N}$, where $\mathbf{w}_N(\bar{\mathbf{A}})$ is the general vector of weights discussed in the text, and subscript N is added to make the population size explicit. Then $\sum_{j=1}^{N} w_{Nj}^{'} = \sum_{j=1}^{N}\left(w_{Nj} - \frac{1}{N}\right) = 0$. Scalar quantity

$$\begin{aligned} x(\bar{\mathbf{A}}, \mathbf{b}, N, n) - \frac{n}{N} &= \left[\mathbf{w}_N(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n) - \frac{n}{N} \\ &= \left(w_{Ni_1} - \frac{1}{N}\right) + \left(w_{Ni_2} - \frac{1}{N}\right) + \cdots + \left(w_{Ni_n} - \frac{1}{N}\right) \\ &= w_{Ni_1}^{'} + w_{Ni_2}^{'} + \cdots + w_{Ni_n}^{'}, \end{aligned}$$

where $1 \leq i_1 < i_2 < \cdots < i_n \leq N$, given a configuration $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$. Thus, by Lemma A.4,

$$\begin{aligned}
G_{\frac{Y\left(\mathbf{w}'_N,N,n\right)}{\sigma\left(\mathbf{w}'_N,N,n\right)}}(t) &= \frac{T\left(\mathbf{w}'_N, N, n, t\right)}{\binom{N}{n}} \\
&= \frac{1}{\binom{N}{n}} \sum_{\forall 1 \le i_1 < i_2 < \cdots < i_n \le N} \mathbb{1}_{w'_{Ni_1} + w'_{Ni_2} + \cdots + w'_{Ni_n} \le t\sigma\left(\mathbf{w}'_N,N,n\right)} \\
&= \frac{1}{\binom{N}{n}} \sum_{\mathbf{b}(N,n) \in \mathcal{B}(N,n)} \mathbb{1}_{x(\bar{\mathbf{A}},\mathbf{b},N,n) - \frac{n}{N} \le t\sigma\left(\mathbf{w}'_N,N,n\right)} \\
&= G_{\frac{X(\bar{\mathbf{A}},N,n) - \frac{n}{N}}{\sigma\left(\mathbf{w}'_N,N,n\right)}}(t),
\end{aligned}$$

so $\lim_{N\to\infty} G_{\frac{X(\bar{\mathbf{A}},N,n)-\frac{n}{N}}{\sigma\left(\mathbf{w}'_N,N,n\right)}}(t) = \Phi(t)$, where $\sigma\left(\mathbf{w}'_N, N, n\right) = \sqrt{\frac{n}{N}\left(1 - \frac{n}{N}\right) \sum_{j=1}^{N} \left(w_{Nj} - \frac{1}{N}\right)^2}$. □

Proof of Theorem 1.12

The proof of Theorem 1.12 makes use of the following result from Höglund (1978), with notation modified for the present work:

Lemma A.5 (Höglund (1978), Main Theorem) *Let $w_1, \ldots, w_N$ be a sequence of real numbers. Let $0 < n < N$ and let $G_Y(t) = \frac{T(\mathbf{w},N,n,t)}{\binom{N}{n}}$, where $T(\mathbf{w}, N, n, t)$ is the total number of sums*

$$y = w_{i_1} + w_{i_2} + \cdots + w_{i_n}, \qquad 1 \le i_1 < i_2 < \cdots < i_n \le N,$$

whose value does not exceed t. Then, for all real t,

$$\left| G_Y(t) - \Phi\left(\frac{t - n\bar{w}}{\left(\frac{n}{N}\left(1 - \frac{n}{N}\right)\sum_{i=1}^{N}\left([\mathbf{w}]_i - \bar{w}\right)^2\right)^{1/2}} \right) \right| \le \frac{C}{\sqrt{\frac{n}{N}\left(1 - \frac{n}{N}\right)}} \frac{\sum_{i=1}^{N}\left|[\mathbf{w}]_i - \bar{w}\right|^3}{\left(\sum_{i=1}^{N}\left([\mathbf{w}]_i - \bar{w}\right)^2\right)^{3/2}},$$

where $\bar{w} = \frac{1}{N}\sum_{i=1}^{N} w_i$.

For the vector of weights, $\mathbf{w}(\bar{\mathbf{A}})$, discussed in the text, $\bar{w} = \frac{1}{N}\sum_{i=1}^{N}[\mathbf{w}(\bar{\mathbf{A}})]_i = \frac{1}{N}$, and

$$G_Y(t) = \frac{T(\mathbf{w}, N, n, t)}{\binom{N}{n}} = \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{x(\bar{\mathbf{A}},\mathbf{b},N,n)\le t} = G_{X(\bar{\mathbf{A}},N,n)}(t).$$

By Lemma A.5, Theorem 1.12 thus follows. □

Proof of Theorem 1.13

The proof of Theorem 1.13 makes use of the following result from Robinson (1978), with notation modified for the present work:

Lemma A.6 (Robinson (1978), Main Theorem) *Let* $\{a_{Ni}\}$ *be a triangular array of real numbers for* $i = 1, \ldots, N$, $N = 2, 3, \ldots$ *and suppose* $\sum_{i=1}^{N} a_{Ni} = 0$, $\sum_{i=1}^{N} a_{Ni}^2 = 1$. *Let* $K_{Nn} = \sum_{i=1}^{n} a_{NR_{Ni}}$, *where* $(R_{N1}, \ldots, R_{NN})$ *is a uniform random permutation of* $(1, \ldots, N)$. *Let* $L_{Nn} = \frac{K_{Nn}}{(\mathrm{Var}\, K_{Nn})^{1/2}}$ *and* $G_{Nn}(t) = Pr(L_{Nn} < t)$. *Set* $p = \frac{n}{N}$ *and* $q = 1 - \frac{n}{N}$. *If condition (c) holds, then*

$$|G_{Nn}(t) - J_{Nn}(t)| < C_4 \times \sum_{i=1}^{N} |a_{Ni}|^5$$

for all t, where C_4 is a function of p only,

$$J_{Nn}(t) = \Phi(t) - H_2(t)\,\phi(t)\,\frac{q-p}{6\,(pq)^{1/2}}\sum_{i=1}^{N} a_{Ni}^3$$
$$- H_3(t)\,\phi(t)\left[\frac{1-6pq}{24pq}\left(\sum_{i=1}^{N} a_{Ni}^4 - 3N^{-1}\right) - \frac{1}{4}N^{-1}\right]$$
$$- H_5(t)\,\phi(t)\,\frac{(q-p)^2}{72pq}\left(\sum_{i=1}^{N} a_{Ni}^3\right)^2,$$

$\phi(t) = \Phi'(t) = (2\pi)^{-1/2} e^{-\frac{1}{2}t^2}$, *and* $H_i(t)\,\phi(t) = (-1)^i\left(\frac{d^i}{dt^i}\right)\phi(t)$. *Condition (c) is as follows:*

Condition (c) *Given $C' > 0$, there exist $\epsilon > 0$, $C > 0$, and $\delta > 0$ not depending on N such that, for any fixed t, the number of indices j, for which $\left|a_{Nj}\,\widehat{x} - t - 2r\pi\right| > \epsilon$, for all $\widehat{x} \in \left(C'\left[\max_i |a_{Ni}|\right]^{-1}, C\left[\sum_{i=1}^{N} |a_{Ni}|^5\right]^{-1}\right)$ and all $r = 0, \pm 1, \pm 2, \ldots$, is greater than δN, for all N.*

Fix N and substitute a_{Ni} with the standardized weight $\widehat{w}_i = \frac{[\mathbf{w}]_i - EW}{\sqrt{N\,\mathrm{Var}\,W}}$, where $EW = \frac{1}{N}\sum_{i=1}^{N}[\mathbf{w}]_i$ and $\mathrm{Var}\,W = \frac{1}{N}\sum_{i=1}^{N}\left([\mathbf{w}]_i - EW\right)^2$. Verify that

$$\sum_{i=1}^{N}\widehat{w}_i = \sum_{i=1}^{N}\frac{w_i - EW}{\sqrt{N\,\mathrm{Var}\,W}} = 0 \quad \text{and} \quad \sum_{i=1}^{N}\widehat{w}_i^2 = \sum_{i=1}^{N}\frac{(w_i - EW)^2}{N\,\mathrm{Var}\,W} = 1.$$

With $B_i \sim \mathrm{Bern}\left(\frac{n}{N}\right)$ and $\mathrm{Var}\,\widehat{W} = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{w}_i - E\widehat{W}\right)^2 = \frac{1}{N}\sum_{i=1}^{N}\widehat{w}_i^2 = \frac{1}{N}$,

$$L_n \equiv \frac{K_n}{(\mathrm{Var}\,K_n)^{1/2}} = \frac{\sum_{i=1}^{n}\widehat{w}_{R_i}}{\left(\mathrm{Var}\sum_{i=1}^{n}\widehat{w}_{R_i}\right)^{1/2}} = \frac{\sum_{i=1}^{n}\frac{(w_{R_i} - EW)}{(N\,\mathrm{Var}\,W)^{1/2}}}{\left(\mathrm{Var}\left(\widehat{w}_1 B_1 + \widehat{w}_2 B_2 + \cdots + \widehat{w}_N B_N\right)\right)^{1/2}}$$
$$= \frac{\sum_{i=1}^{n} w_{R_i} - nEW}{(N\,\mathrm{Var}\,W)^{1/2}\left(\frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}N\,\mathrm{Var}\,\widehat{W}\right)^{1/2}}$$
$$= \frac{X(\bar{\mathbf{A}}, N, n) - EX(\bar{\mathbf{A}}, N, n)}{\left(\mathrm{Var}\,X(\bar{\mathbf{A}}, N, n)\right)^{1/2}},$$

and Theorem 1.13 follows. □

Proof of Theorem 1.14

When k_ω agents have the same non-zero weight ω, and all other $N - k_\omega$ agents have zero weight, $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is non-zero only for integer multiples of $\omega : i\omega$. We now determine the allowable values of i. Given N, n, k_ω, the smallest possible value of i is $\max\{0, n - (N - k_\omega)\}$. If $N - k_\omega \geq n$, that is, if the number of agents with zero weight is greater than or equal to n, then there exists at least one configuration $\mathbf{b}(N,n)$ for which $x(\bar{\mathbf{A}}, \mathbf{b}, N, n) = 0$ and $g_{X(\bar{\mathbf{A}},N,n)}(0) > 0$. If $N - k_\omega < n$, so that the number of agents with zero weight is less than n, then $\min \operatorname{supp} X(\bar{\mathbf{A}}, N, n) = [n - (N - k_\omega)] \times \omega$. Given N, n, k_ω, the largest possible value of i is $\min\{n, k_\omega\}$. If $n \leq k_\omega$, then there exists at least one configuration $\mathbf{b}(N,n)$ for which $x(\bar{\mathbf{A}}, \mathbf{b}, N, n) = n\omega$ and $g_{X(\bar{\mathbf{A}},N,n)}(n\omega) > 0$. If $n > k_\omega$, then $\max \operatorname{supp} X(\bar{\mathbf{A}}, N, n) = k_\omega \omega$. Therefore, we define the set

$$\mathcal{I} = \{\max\{0, n - (N - k_\omega)\}, \max\{0, n - (N - k_\omega)\} + 1, \ldots, \min\{n, k_\omega\}\}.$$

For all values $i \in \mathcal{I}$,

$$g_{X(\bar{\mathbf{A}},N,n)}(i\omega) = \frac{\binom{k_\omega}{i}\binom{N-k_\omega}{n-i}}{\binom{N}{n}},$$

which is hypergeometric. For all values $i \notin \mathcal{I}$, $g_{X(\bar{\mathbf{A}},N,n)}(i\omega) = 0$. $\square$

Proof of Theorem 1.15

For all $i, \ell \in \{1, \ldots, N\}$, $[\bar{\mathbf{A}}]_{i\ell} \in \left\{0, \frac{1}{k}\right\}$. Then $\mathbf{d}_w^-(\bar{\mathbf{A}}) = \frac{1}{N}\bar{\mathbf{A}}^T\mathbf{1} = \frac{1}{N}\frac{1}{k}\mathbf{d}^-(\mathbf{A})$ and $D_w^-(\bar{\mathbf{A}}) = \frac{1}{Nk}D^-(\mathbf{A})$. With

$$\begin{aligned}\operatorname{Var} D_w^-(\bar{\mathbf{A}}) = \frac{1}{N}\sum_{i=1}^{N}\left(\left[\mathbf{d}_w^-(\bar{\mathbf{A}})\right]_i - \frac{1}{N}\right)^2 &= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{Nk}\left[\mathbf{d}^-(\mathbf{A})\right]_i - \frac{k}{Nk}\right)^2 \\ &= \left(\frac{1}{Nk}\right)^2 \operatorname{Var} D^-(\mathbf{A}),\end{aligned}$$

it follows from Theorem 1.9 that

$$\operatorname{Var}\widehat{F}_{avg}\left(\bar{\mathbf{A}},N,n\right)=\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}N\left(\frac{1}{Nk}\right)^{2}\operatorname{Var}D^{-}\left(\mathbf{A}\right).$$

Construct the ordered multiset $\{w_s\}_{s=1}^{N}$ from the elements of $\mathbf{d}_w^-\left(\bar{\mathbf{A}}\right)$ so that $w_s \leq w_{s'}$ whenever $s \leq s'$. From Theorem 1.10, $\min\operatorname{supp}\widehat{F}_{avg}\left(\bar{\mathbf{A}},N,n\right)=\sum_{s=1}^{n}w_s=\frac{1}{Nk}\sum_{s=1}^{n}v_s$ and $\max\operatorname{supp}\widehat{F}_{avg}\left(\bar{\mathbf{A}},N,n\right)=\sum_{s=N-n+1}^{N}w_s=\frac{1}{Nk}\sum_{s=N-n+1}^{n}v_s$. $\square$

Proof of Theorem 1.16

$X\left(\bar{\mathbf{A}},N,n\right)=\frac{\widetilde{W}_1(\bar{\mathbf{A}})+\cdots+\widetilde{W}_n(\bar{\mathbf{A}})}{n}\times f$, with $f=\frac{n}{N}$ fixed as $N\to\infty$, so $\frac{\widetilde{W}_1(\bar{\mathbf{A}})+\cdots+\widetilde{W}_n(\bar{\mathbf{A}})}{n}=\frac{1}{f}X\left(\bar{\mathbf{A}},N,n\right)$. Suppose that $\operatorname{Var}\widetilde{W}_i\left(\bar{\mathbf{A}}\right)=N^2\operatorname{Var}W\left(\bar{\mathbf{A}}\right)$ is finite. Then by the Classical Central Limit Theorem,

$$n^{1/2}\left(\frac{1}{f}X\left(\bar{\mathbf{A}},N,n\right)-E\widetilde{W}_i\left(\bar{\mathbf{A}}\right)\right)\xrightarrow[d]{}\mathcal{N}\left(0,\operatorname{Var}\widetilde{W}_i\left(\bar{\mathbf{A}}\right)\right),$$

where $E\widetilde{W}_i\left(\bar{\mathbf{A}}\right)=1$. Now suppose that $\operatorname{Var}\widetilde{W}_i\left(\bar{\mathbf{A}}\right)$ is infinite and $E\widetilde{W}_i\left(\bar{\mathbf{A}}\right)=1$. Specifically, $\Pr\left[\widetilde{W}_i\left(\bar{\mathbf{A}}\right)>t\right]\sim L(t)\,t^{-\xi}$, where $L(t)$ is a slowly varying function and $\xi\in(1,2)$ since $\operatorname{Var}\widetilde{W}_i\left(\bar{\mathbf{A}}\right)$ is infinite and $E\widetilde{W}_i\left(\bar{\mathbf{A}}\right)$ is finite. As discussed in Nolan (2014), a specific case of the Generalized Central Limit Theorem is as follows, with notation adapted for the present setting:

Lemma A.7 *Let $\widetilde{W}_1\left(\bar{\mathbf{A}}\right),\widetilde{W}_2\left(\bar{\mathbf{A}}\right),\ldots$ be independent, identically distributed random variables. Suppose that $\Pr\left[\widetilde{W}_i\left(\bar{\mathbf{A}}\right)>t\right]\sim C^+t^{-\xi}$ and $\Pr\left[\widetilde{W}_i\left(\bar{\mathbf{A}}\right)<-t\right]\sim C^-\left|t\right|^{-\xi}$ as $t\to\infty$ with $1<\xi<2$ and $C^++C^->0$. Set $\beta=\frac{C^+-C^-}{C^++C^-}$. Then*

$$\frac{\widetilde{W}_1\left(\bar{\mathbf{A}}\right)+\cdots+\widetilde{W}_n\left(\bar{\mathbf{A}}\right)-nE\widetilde{W}_i\left(\bar{\mathbf{A}}\right)}{n^{1/\xi}}\xrightarrow[d]{}\widetilde{S}\left(\xi,\beta,\widetilde{\gamma},0;1\right),$$

where $\widetilde{S}\left(\xi,\beta,\widetilde{\gamma},0;1\right)$ is a stable distribution with characteristic function

$$E\exp\left(iu\widetilde{W}_i\left(\bar{\mathbf{A}}\right)\right)=\exp\left(-\widetilde{\gamma}^{\xi}\left|u\right|^{\xi}\left[1-i\beta\left(\tan\frac{\pi\xi}{2}\right)\times\operatorname{sign}u\right]\right)$$

when $\xi \neq 1$.

With $\frac{\widetilde{W}_1(\bar{\mathbf{A}})+\cdots+\widetilde{W}_n(\bar{\mathbf{A}})-nE\widetilde{W}_i(\bar{\mathbf{A}})}{n^{1/\xi}} = n^{1-1/\xi}\left(\frac{\widetilde{W}_1(\bar{\mathbf{A}})+\cdots+\widetilde{W}_n(\bar{\mathbf{A}})}{n} - E\widetilde{W}_i\left(\bar{\mathbf{A}}\right)\right)$, we obtain the result

$$n^{1-1/\xi}\left(\frac{1}{f}X\left(\bar{\mathbf{A}},N,n\right) - E\widetilde{W}_i\left(\bar{\mathbf{A}}\right)\right) \xrightarrow[d]{} \widetilde{S}\left(\xi,\beta,\widetilde{\gamma},0;1\right). \quad \square$$

Proof of Theorem 1.17

$$EX\left(\bar{\mathbf{A}},N,n,(\gamma_i)_{i=1}^N\right) = E\left[w_1\left(\bar{\mathbf{A}}\right)B_1 + w_2\left(\bar{\mathbf{A}}\right)B_2 + \cdots + w_N\left(\bar{\mathbf{A}}\right)B_N\right] = \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T\boldsymbol{\mu},$$

and

$$\begin{aligned}\operatorname{Var}X\left(\bar{\mathbf{A}},N,n,(\gamma_i)_{i=1}^N\right) &= \operatorname{Var}\left[w_1\left(\bar{\mathbf{A}}\right)B_1 + w_2\left(\bar{\mathbf{A}}\right)B_2 + \cdots + w_N\left(\bar{\mathbf{A}}\right)B_N\right] \\ &= \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]^T\boldsymbol{\Sigma}\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right],\end{aligned}$$

with $\boldsymbol{\mu} = E\mathbf{B}$ as the conditional mean vector for $\mathbf{B}$ and $\boldsymbol{\Sigma}$ as the $N\times N$ conditional covariance matrix for $\mathbf{B}$. The $N\times 1$ random vector $\mathbf{B}\equiv\mathbf{B}\left(N,n,\mathbf{s},\boldsymbol{\psi}\right)$ is distributed according to Fisher's multivariate non-central hypergeometric distribution (see McCullagh and Nelder (1989)). Each agent exists in the population with frequency 1, so define the $N\times 1$ frequency vector $\mathbf{s} = \mathbf{1}$. Next, let $\boldsymbol{\phi}$ be an $N\times 1$ vector of probabilities whose elements are defined as $\phi_i = \Pr\left[B_i = 1|\,\gamma_i\right]$; construct the $N\times 1$ vector $\boldsymbol{\psi}$ with element $\psi_i = \frac{\phi_i}{1-\phi_i}/\frac{\phi_k}{1-\phi_k}$ relative to some agent k and $\psi_k\equiv 1$. $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be approximated by solving the following system of equations:

$$\begin{aligned}&\sum_{i=1}^N \mu_i = n,\\ &\psi_j = \frac{\mu_j\left(s_k-\mu_k\right)-\Sigma_{jk}}{\left(s_j-\mu_j\right)\mu_k-\Sigma_{jk}},\quad \forall j\in\{1,\ldots,N\}\setminus\{k\},\text{ and}\\ &\boldsymbol{\Sigma} = \frac{N}{N-1}\left(\operatorname{diag}\boldsymbol{\zeta} - \frac{\boldsymbol{\zeta}\boldsymbol{\zeta}^T}{\mathbf{1}^T\boldsymbol{\zeta}}\right),\quad\text{with}\quad \frac{1}{\zeta_j} = \frac{1}{\mu_j}+\frac{1}{s_j-\mu_j}.\end{aligned}$$

Now, the number of equations in the system can be reduced by noting that $\mu_i = \mu_j$ when $\phi_i = \phi_j$, so $\zeta_i = \zeta_j$, $\psi_i = \psi_j$, and $\Sigma_{ik} = \Sigma_{jk}$. Therefore, partition agent indices into Θ categories according to their conditional probabilities, that is, agents i, j are in category θ if $\phi_i = \phi_j = \rho_\theta$. Define the odds ratio for agents in category θ relative to category k as: $\widehat{\psi}_\theta = \frac{\rho_\theta}{1-\rho_\theta} / \frac{\rho_k}{1-\rho_k}$, with $\widehat{\psi}_k \equiv 1$. Define the $\Theta \times 1$ vector $\widehat{\boldsymbol{\mu}}$ across the Θ categories, setting $\mu_i = \widehat{\mu}_\theta$ for each agent i from category θ, and setting $\zeta_i = \widehat{\mu}_\theta (1 - \widehat{\mu}_\theta)$ for each agent i from category θ. Define the $\Theta \times \Theta$ matrix $\widehat{\boldsymbol{\Sigma}}$ with element $\widehat{\Sigma}_{\theta k}$ equal to the conditional covariance $\operatorname{Cov}\left(B_i, B_j\right)$ between agent i in category θ and agent j in category k. The system of equations then collapses to the following:

$$\sum_{\theta=1}^{\Theta} \sum_{\substack{i \in \{1,\ldots,N\} \\ \text{s.t. } \phi_i = \rho_\theta}} \widehat{\mu}_\theta = n,$$

$$\widehat{\psi}_\theta = \frac{\widehat{\mu}_\theta (1 - \widehat{\mu}_k) - \widehat{\Sigma}_{\theta k}}{(1 - \widehat{\mu}_\theta) \widehat{\mu}_k - \widehat{\Sigma}_{\theta k}}, \quad \forall \theta \in \{1, \ldots, \Theta\} \setminus \{k\}, \text{ and}$$

$$\boldsymbol{\Sigma} = \frac{N}{N-1} \left(\operatorname{diag} \boldsymbol{\zeta} - \frac{\boldsymbol{\zeta}\boldsymbol{\zeta}^T}{\mathbf{1}^T \boldsymbol{\zeta}} \right). \quad \square$$

Proof of Theorem A.1

The first three lines of this proof have a similar (but not equal) construction to those in the proof for Theorem 1.2 in Seneta (1981), so they are accordingly marked. Define the resolvent $\mathbf{R}(z) = (\mathbf{I} - z\bar{\mathbf{A}})^{-1}$ with matrix element $[\mathbf{R}(z)]_{ij}$ and $z \neq \lambda_i^{-1}$ for every eigenvalue λ_i of $\bar{\mathbf{A}}$, $i \in \{1, \ldots, t\}$. Eigenvalues λ_i are ordered by weakly descending modulus with m_i, the algebraic multiplicity of λ_i, greater than m_{i+1} whenever $|\lambda_i| = |\lambda_{i+1}|$.

$$\mathbf{R}(z) = \frac{\operatorname{Adj}(\mathbf{I} - z\bar{\mathbf{A}})}{\det(\mathbf{I} - z\bar{\mathbf{A}})},$$

where $\mathrm{Adj}\,(\mathbf{X})$ is the adjugate matrix of $\mathbf{X}$, or the transpose of the cofactor matrix of $\mathbf{X}$.

$$[\mathbf{R}\,(z)]_{ij} = \frac{c_{ij}\,(z)}{(1-z)\,(1-z\lambda_2)^{m_2}\cdots(1-z\lambda_t)^{m_t}}, \tag{A.2}$$

where $c_{ij}\,(z)$ is a polynomial in z of degree at most $N-1$. By partial fraction decomposition,

$$[\mathbf{R}(z)]_{ij} = \eta_{ij}\,(z) + \frac{\gamma_{ij}}{1-z} + \sum_{s=0}^{m_2-1} \frac{\beta_{ij,m_2-s}}{(1-z\lambda_2)^{m_2-s}} + \sum_{s=0}^{m_3-1} \frac{\beta_{ij,m_3-s}}{(1-z\lambda_3)^{m_3-s}} + \cdots, \tag{A.3}$$

where γ_{ij}, $\{\beta_{ij,m_2-s}\}_{s=0}^{m_2-1}$, $\{\beta_{ij,m_3-s}\}_{s=0}^{m_3-1}$, ... are constants, and $\eta_{ij}\,(z) = \eta_{ij,0} + \eta_{ij,1}\,z + \cdots + \eta_{ij,N-2}\,z^{N-2}$ is a polynomial of degree at most $N-2$. For $|z| < \frac{1}{\lambda_1} = 1$,

$$[\mathbf{R}\,(z)]_{ij} = \eta_{ij}\,(z) + \gamma_{ij} \sum_{q=0}^{\infty} z^q + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \left[\sum_{q=0}^{\infty} \binom{-m_2+s}{q} (-z\lambda_2)^q \right] + \cdots. \tag{A.4}$$

Now, since $s \in \{0,1,\ldots,m_2-1\}$, $-m_2+s<0$ so $\binom{-m_2+s}{q} = (-1)^q \binom{m_2-s+q-1}{q}$ and

$$\begin{aligned}
[\mathbf{R}\,(z)]_{ij} &= \sum_{q=0}^{\infty} z^q\,[\bar{\mathbf{A}}^q]_{ij} \\
&= \eta_{ij}\,(z) + \gamma_{ij} \sum_{q=0}^{\infty} z^q + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \left[\sum_{q=0}^{\infty} \binom{m_2-s+q-1}{q} (z\lambda_2)^q \right] + \cdots.
\end{aligned}$$

Matching coefficients of z^q,

$$[\bar{\mathbf{A}}^q]_{ij} = \begin{cases} \eta_{ij,q} + \left[\mathbf{w}_\infty^T\right]_j + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \binom{m_2-s+q-1}{q} \lambda_2^q + \cdots & \text{for } 1 \le q \le N-2 \\ \left[\mathbf{w}_\infty^T\right]_j + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \binom{m_2-s+q-1}{q} \lambda_2^q + \cdots & \text{for } q > N-2 \end{cases},$$

with $\eta_{ij,q} = 0$ for $q > N-2$ and $\gamma_{ij} = \left[\mathbf{w}_\infty^T\right]_j$ by Lemma A.2. It follows that

$$\begin{aligned}
\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) &= \sum_{j=1}^{N} \left[\bar{\mathbf{A}}^q\right]_{ij} [\mathbf{b}]_j \\
&= \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) + \sum_{j=1}^{N} \left[\eta_{ij,q} + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \binom{m_2-s+q-1}{q} \lambda_2^q\right] [\mathbf{b}]_j \\
&\qquad + \sum_{j=1}^{N} \left[\sum_{s=0}^{m_3-1} \beta_{ij,m_3-s} \binom{m_3-s+q-1}{q} \lambda_3^q\right] [\mathbf{b}]_j \\
&= \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) + \sum_{j=1}^{N} \left[\eta_{ij,q} + \sum_{s=0}^{m_2-1} \beta_{ij,m_2-s} \binom{m_2-s+q-1}{q} \lambda_2^q\right] [\mathbf{b}]_j \\
&\qquad + \mathcal{O}\left((q+m_3-1)^{m_3-1} |\lambda_3|^q\right)
\end{aligned}$$

since

$$\begin{aligned}
\binom{m_3-s+q-1}{q} &= \frac{(q+1)(q+2)\cdots(q+(m_3-s-1))}{(m_3-s-1)!} \\
&< \frac{(q+(m_3-s-1))^{m_3-s-1}}{(m_3-s-1)!} \\
&\leq \frac{(q+m_3-1)^{m_3-1}}{(m_3-s-1)!}.
\end{aligned}$$

When $m_2 = 1$,

$$\begin{aligned}
\widehat{f}_i^{(q)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) = \widehat{f}^{(\infty)}\left(\bar{\mathbf{A}}, \mathbf{b}, N, n\right) + \sum_{j=1}^{N} \left[\eta_{ij,q} + \lambda_2^q \beta_{ij,m_2-0}\right] [\mathbf{b}]_j \\
+ \mathcal{O}\left((q+m_3-1)^{m_3-1} |\lambda_3|^q\right).
\end{aligned}$$

To determine the rate of convergence, note that

$$\binom{m_2-s+q-1}{q} \in \left(\frac{q^{m_2-s-1}}{(m_2-s-1)!}, \frac{(q+(m_2-s-1))^{m_2-s-1}}{(m_2-s-1)!}\right),$$

$$q^{m_2-s-1} = q^{m_2-1} q^{-s},$$

and

$$
\begin{aligned}
&((q+1)+(m_2-s-1))^{m_2-s-1} \\
&\qquad = ((q+1)+(m_2-s-1))^{m_2-1}((q+1)+(m_2-s-1))^{-s} \\
&\qquad \leq ((q+1)+(m_2-1))^{m_2-1}((q+1)+(m_2-s-1))^{-s}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\left|\frac{\widehat{f}_i^{(q+1)}(\bar{\mathbf{A}},\mathbf{b},N,n)-\widehat{f}^{(\infty)}(\bar{\mathbf{A}},\mathbf{b},N,n)}{\widehat{f}_i^{(q)}(\bar{\mathbf{A}},\mathbf{b},N,n)-\widehat{f}^{(\infty)}(\bar{\mathbf{A}},\mathbf{b},N,n)}\right| \\
&\quad = \left|\frac{\sum_{j=1}^{N}\left[\eta_{ij,q+1}+\sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q}{q+1}\lambda_2^{q+1}+\cdots\right][\mathbf{b}]_j}{\sum_{j=1}^{N}\left[\eta_{ij,q}+\sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^{q}+\cdots\right][\mathbf{b}]_j}\right| \\
&\quad \leq \left|\frac{\sum_{j=1}^{N}\left[\eta_{ij,q+1}+\sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\frac{((q+1)+(m_2-s-1))^{m_2-s-1}}{(m_2-s-1)!}\lambda_2^{q+1}+\cdots\right][\mathbf{b}]_j}{\sum_{j=1}^{N}\left[\eta_{ij,q}+\sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\frac{q^{m_2-s-1}}{(m_2-s-1)!}\lambda_2^{q}+\cdots\right][\mathbf{b}]_j}\right| \\
&\quad = \mathcal{O}\left(\left(1+\frac{m_2}{q}\right)^{m_2-1}|\lambda_2|\right).
\end{aligned}
$$

As $q\to\infty$, by Lemma A.2,

$$
\widehat{f}_i^{(q)}(\bar{\mathbf{A}},\mathbf{b},N,n)=\widehat{f}^{(\infty)}(\bar{\mathbf{A}},\mathbf{b},N,n)+\mathcal{O}\left(q^{m_2-1}|\lambda_2|^q\right). \quad \square
$$

Proof of Theorem A.2

Demonstrating that $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is always symmetric if $g_{W(\bar{\mathbf{A}})}(t)$ is symmetric:

Start with the set $\mathcal{B}(N,n)$ of configurations. Remove all configurations $\mathbf{b}(N,n)$ for which $[\mathbf{w}(\bar{\mathbf{A}})]^T\mathbf{b}(N,n)=\frac{n}{N}$. Partition the set

$$
\mathcal{B}(N,n)\setminus\left\{\mathbf{b}(N,n)\in\mathcal{B}(N,n):[\mathbf{w}(\bar{\mathbf{A}})]^T\mathbf{b}(N,n)=\frac{n}{N}\right\}
$$

into pairs $(\mathbf{b}(N,n),\mathbf{b}'(N,n))$. Choose $\mathbf{b}'(N,n)$ to pair with $\mathbf{b}(N,n)$ according to the following procedure:

For each index i for which $[\mathbf{b}]_i = 1$, select a unique index j for which $w_i - \frac{1}{N} = \frac{1}{N} - w_j$. Accordingly, set $[\mathbf{b}']_j = 1$. Then go to the next index i for which $[\mathbf{b}]_i = 1$, choose a different index j for which $w_i - \frac{1}{N} = \frac{1}{N} - w_j$, and set $[\mathbf{b}']_j = 1$. Continue until $\mathbf{b}'(N,n)$ has been constructed. The construction of $\mathbf{b}'(N,n)$ for every pair $(\mathbf{b}(N,n), \mathbf{b}'(N,n))$ is guaranteed because $g_{W(\bar{\mathbf{A}})}(t)$ is symmetric, so multiset $\{w_i\}_{i=1}^{N} \setminus \left\{w_i : w_i = \frac{1}{N}\right\}$ can be partitioned into pairs such that $w_i - \frac{1}{N} = \frac{1}{N} - w_j$ for each pair (w_i, w_j). It then follows that, for every pair $(\mathbf{b}(N,n), \mathbf{b}'(N,n))$,

$$\begin{aligned}
&\left(x(\bar{\mathbf{A}}, \mathbf{b}, N, n) - \frac{n}{N}\right) + \left(x(\bar{\mathbf{A}}, \mathbf{b}, N, n) - \frac{n}{N}\right) \\
&\qquad = \left([\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N,n) - \frac{n}{N}\right) + \left([\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}'(N,n) - \frac{n}{N}\right) \\
&\qquad = \left(\sum_{\substack{i \in \{1,\ldots,N\} \\ \text{s.t. } b_i = 1}} w_i - \frac{n}{N}\right) + \left(\sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } b_j = 1}} w_j - \frac{n}{N}\right) \\
&\qquad = \left(\sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } b_j = 1}} \left(\frac{2}{N} - w_j\right) - \frac{n}{N}\right) + \left(\sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } b_j = 1}} w_j - \frac{n}{N}\right) \\
&\qquad = \frac{2n}{N} - \frac{2n}{N} + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } b_j = 1}} w_j - \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } b_j = 1}} w_j \\
&\qquad = 0,
\end{aligned}$$

so $g_{X(\bar{\mathbf{A}},N,n)}(t)$ must be symmetric.

Demonstrating that $g_{X(\bar{\mathbf{A}},N,n)}(t)$ is always symmetric when $f = 0.5$:

Let $f = \frac{n}{N} = 0.5$ $(0.5N \in \mathbb{Z}_+)$. Consider all configurations $\mathbf{b}(N, 0.5N) \in \mathcal{B}(N, 0.5N)$ for which $[\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N, 0.5N) \neq \frac{n}{N}$. $g_{X(\bar{\mathbf{A}},N,0.5N)}(t)$ is symmetric if and only if

$$[\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N, 0.5N) - \frac{n}{N} = \frac{n}{N} - [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}'(N, 0.5N)$$

for every pair $(\mathbf{b}(N, 0.5N), \mathbf{b}'(N, 0.5N))$, that is, if and only if,

$$[\mathbf{w}(\bar{\mathbf{A}})]^T (\mathbf{b}(N, 0.5N) + \mathbf{b}'(N, 0.5N)) = 0.5 + 0.5 = 1$$

for every pair $(\mathbf{b}(N, 0.5N), \mathbf{b}'(N, 0.5N))$. Set $\mathbf{b}'(N, 0.5N) = \mathbf{1} - \mathbf{b}(N, 0.5N) \in \mathcal{B}(N, 0.5N)$, and it follows that $g_{X(\bar{\mathbf{A}},N,0.5N)}(t)$ is symmetric.

Demonstrating that $g_{X(\bar{\mathbf{A}},N,n)}(t) = g_{X(\bar{\mathbf{A}},N,N-n)}(1-t)$, so that $\text{Var}\, X(\bar{\mathbf{A}}, N, n) = \text{Var}\, X(\bar{\mathbf{A}}, N, N-n)$, $\text{Skew}\, X(\bar{\mathbf{A}}, N, n) = -\text{Skew}\, X(\bar{\mathbf{A}}, N, N-n)$, and $\text{Kurt}\, X(\bar{\mathbf{A}}, N, n) = \text{Kurt}\, X(\bar{\mathbf{A}}, N, N-n)$:

$g_{X(\bar{\mathbf{A}},N,n)}(t) = g_{X(\bar{\mathbf{A}},N,N-n)}(1-t)$ if and only if there exist pairs $(\mathbf{b}(N, n), \mathbf{b}(N, N-n))$ of configurations such that

$$[\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N, n) = 1 - [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{b}(N, N-n)$$

for every pair, that is, if and only if

$$[\mathbf{w}(\bar{\mathbf{A}})]^T (\mathbf{b}(N, n) + \mathbf{b}(N, N-n)) = 1$$

for every pair. Set $\mathbf{b}(N, n) = \mathbf{1} - \mathbf{b}(N, N-n)$, with $\mathbf{b} \in \mathcal{B}(N, n)$ and $\mathbf{b}(N, N-n) \in \mathcal{B}(N, N-n)$, so $g_{X(\bar{\mathbf{A}},N,n)}(t) = g_{X(\bar{\mathbf{A}},N,N-n)}(1-t)$. It then follows that $\text{Var}\, X(\bar{\mathbf{A}}, N, n) = \text{Var}\, X(\bar{\mathbf{A}}, N, N-n)$, $\text{Skew}\, X(\bar{\mathbf{A}}, N, n) = -\text{Skew}\, X(\bar{\mathbf{A}}, N, N-n)$, and $\text{Kurt}\, X(\bar{\mathbf{A}}, N, n) = \text{Kurt}\, X(\bar{\mathbf{A}}, N, N-n)$. $\square$

Proof of Theorem A.3

Construct the ordered multiset $\{w_s\}_{s=1}^{N}$ from the elements of $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ so that $w_s \leq w_{s'}$ whenever $s \leq s'$. For $j \in \{-2, -1, 0, 1, 2, \}$, define

$$\begin{aligned}\ell_{n+j} &= \max \operatorname{supp} X\left(\bar{\mathbf{A}}, N, n+j\right) - \min \operatorname{supp} X\left(\bar{\mathbf{A}}, N, n+j\right) \\ &= \sum_{s=N-n+1-j}^{N} w_s - \sum_{s=1}^{n+j} w_s.\end{aligned}$$

First, $\ell_n - \ell_{n-1} = w_{N-n+1} - w_n \geq 0$ if and only if $w_{N-n+1} \geq w_n$, that is, if and only if

$$n \leq N - n + 1 \Leftrightarrow 2n \leq N + 1 \Leftrightarrow 0 \leq f \leq \frac{1}{2}\left(\frac{N+1}{N}\right) \Leftrightarrow 0 \leq f \leq \frac{1}{2},$$

so the width of $\operatorname{supp} X\left(\bar{\mathbf{A}}, N, n\right)$ weakly increases for $f \in \left[0, \frac{1}{2}\right]$.

Next, $\ell_{n+1} - \ell_n = w_{N-n} - w_{n+1} \leq 0$ if and only if $w_{N-n} \leq w_{n+1}$, that is, if and only if

$$n + 1 \geq N - n \Leftrightarrow 2n \geq N - 1 \Leftrightarrow 1 \geq f \geq \frac{1}{2}\left(\frac{N-1}{N}\right) \Leftrightarrow 1 \leq f \leq \frac{1}{2},$$

so the width of $\operatorname{supp} X\left(\bar{\mathbf{A}}, N, n\right)$ weakly decreases for $f \in \left[\frac{1}{2}, 1\right]$.

Last,

$$\begin{aligned}(\ell_n - \ell_{n-1}) - (\ell_{n-1} - \ell_{n-2}) &= (w_{N-n+1} - w_n) - (w_{N-n+2} - w_{n-1}) \\ &= (w_{N-n+1} - w_{N-n+2}) + (w_{n-1} - w_n) \leq 0,\end{aligned}$$

so the width of $\operatorname{supp} X\left(\bar{\mathbf{A}}, N, n\right)$ weakly increases at a weakly decreasing rate for $f \in \left[0, \frac{1}{2}\right]$, and

$$\begin{aligned}(\ell_{n+2} - \ell_{n+1}) - (\ell_{n+1} - \ell_n) &= (w_{N-n-1} - w_{n+2}) - (w_{N-n} - w_{n+1}) \\ &= (w_{N-n-1} - w_{N-n}) + (w_{n+1} - w_{n+2}) \leq 0,\end{aligned}$$

so the width of supp $X(\bar{\mathbf{A}}, N, n)$ weakly decreases at a weakly increasing rate for $f \in \left[\frac{1}{2}, 1\right]$. $\square$

Proof of Theorem A.4

$\mathrm{Cov}\left(X_1\left(\bar{\mathbf{A}}, N, n\right), X_2\left(\bar{\mathbf{A}}, N, n\right)\right) = \mathrm{Cov}\left[\mathbf{w}_1^T \mathbf{B}\left(N, n\right), \mathbf{w}_2^T \mathbf{B}\left(N, n\right)\right]$, where $\mathbf{B}(N, n)$ is a random vector whose elements are $B_i \sim \mathrm{Bern}\,\frac{n}{N}$, for all $i \in \{1, \ldots, N\}$. Therefore,

$$
\begin{aligned}
&\mathrm{Cov}\left(X_1\left(\bar{\mathbf{A}}, N, n\right), X_2\left(\bar{\mathbf{A}}, N, n\right)\right) \\
&\quad = \mathrm{Cov}\left(w_{11}B_1 + w_{12}B_2 + \cdots + w_{1N}B_N,\ w_{21}B_1 + w_{22}B_2 + \cdots + w_{2N}B_N\right) \\
&\quad = \sum_{i=1}^{N} \mathrm{Cov}\left(w_{1i}B_i,\ w_{2i}B_i\right) + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathrm{Cov}\left(w_{1i}B_i,\ w_{2j}B_j\right) \\
&\quad = \left(\mathrm{Var}\,B_i\right) \sum_{i=1}^{N} w_{1i}w_{2i} + \left(E\left[B_iB_j\right] - \left(EB_i\right)\left(EB_j\right)\right) \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{1i}w_{2j} \\
&\quad = \frac{n}{N}\left(1 - \frac{n}{N}\right) \sum_{i=1}^{N} w_{1i}w_{2i} - \frac{n\left(1 - \frac{n}{N}\right)}{N\left(N-1\right)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{1i}w_{2j} \\
&\quad = \frac{n}{N}\left(1 - \frac{n}{N}\right) \times \left(\sum_{i=1}^{N} w_{1i}w_{2i} - \frac{1}{N-1} \sum_{k=1} N \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{1i}w_{2j}\right).
\end{aligned}
$$

Now,

$$\begin{aligned}
\mathrm{Cov}\left(w_1, w_2\right) &= \frac{1}{N}\sum_{i=1}^{N}\left(w_{1i} - \frac{\sum_{j=1}^{N} w_{1j}}{N}\right)\left(w_{2i} - \frac{\sum_{j=1}^{N} w_{2j}}{N}\right) \\
&= \frac{1}{N}\left(\sum_{i=1}^{N} w_{1i}w_{2i} - \frac{1}{N}\sum_{i=1}^{N} w_{1i}\sum_{i=1}^{N} w_{2i}\right) \\
&= \frac{1}{N}\left(\sum_{i=1}^{N} w_{1i}w_{2i} - \frac{1}{N}\sum_{i=1}^{N} w_{1i}w_{2i} - \frac{1}{N}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} w_{1i}w_{2j}\right) \\
&= \frac{1}{N}\frac{N-1}{N}\left(\sum_{i=1}^{N} w_{1i}w_{2i} - \frac{1}{N-1}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} w_{1i}w_{2j}\right),
\end{aligned}$$

so $\mathrm{Cov}\left(X_1\left(\bar{\mathbf{A}}, N, n\right), X_2\left(\bar{\mathbf{A}}, N, n\right)\right) = \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N\,\mathrm{Cov}\left(W_1, W_2\right)\right)$. □

Proof of Theorem A.5

By Corollary 1.2, $E\,\widehat{\mathbf{F}}^{(q)}\left(\bar{\mathbf{A}}, N, n\right) = \frac{n}{N}\mathbf{1}$. Next,

$$\begin{aligned}
\left[\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}}, N, n\right)\right]_{ik} &= \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}\left(N\,\mathrm{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right)\right) \\
&= \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N}{N-1}N\left(\frac{1}{N}\sum_{j=1}^{N}\left(\left[\mathbf{w}_{i,a}^{(q)}\right]_j - \frac{1}{N}\right)\left(\left[\mathbf{w}_{k,a}^{(q)}\right]_j - \frac{1}{N}\right)\right)
\end{aligned}$$

by Theorem A.4, with $\left[\mathbf{w}_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right]^T = [\bar{\mathbf{A}}^q]_{i*}$ and $\left[\mathbf{w}_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right]^T = [\bar{\mathbf{A}}^q]_{k*}$. From Theorem A.1,

$$\begin{aligned}\operatorname{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right) = \frac{1}{N}\sum_{j=1}^{N}\Bigg(&\eta_{ij,q} + \sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q \\ &+\mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right) + \left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\Bigg) \\ &\times\left(\eta_{kj,q} + \sum_{s=0}^{m_2-1}\beta_{kj,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q\right. \\ &\left.+\mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right) + \left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\right).\end{aligned}$$

As $q\to\infty$, $\eta_{ij,q} = \eta_{kj,q} = 0$ and $\binom{m_2-s+q-1}{q} \sim \frac{q^{m_2-s-1}}{(m_2-s-1)!} < \frac{q^{m_2-1}}{(m_2-s-1)!}$, so

$$\begin{aligned}\operatorname{Cov}\left(W_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right), W_{k,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right) = \frac{1}{N}\sum_{j=1}^{N}\Bigg(&\sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q \\ &+\mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right)\Bigg) \\ \times\left(\sum_{r=0}^{m_2-1}\beta_{kj,m_2-r}\binom{m_2-r+q-1}{q}\lambda_2^q + \mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right)\right)& \\ +\frac{1}{N}\sum_{j=1}^{N}\Bigg(&\sum_{s=0}^{m_2-1}\beta_{ij,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q \\ &+\mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right)\Bigg)\left(\left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\right) \\ +\frac{1}{N}\sum_{j=1}^{N}\Bigg(&\sum_{s=0}^{m_2-1}\beta_{kj,m_2-s}\binom{m_2-s+q-1}{q}\lambda_2^q \\ +\mathcal{O}\left((q+m_3-1)^{m_3-1}|\lambda_3|^q\right)\Bigg)&\left(\left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\right) + \frac{1}{N}\sum_{j=1}^{N}\left(\left[\mathbf{w}_\infty^T\right]_j - \frac{1}{N}\right)^2 \\ &= \mathcal{O}\left(q^{2m_2-2}|\lambda_2|^{2q}\right) + \operatorname{Var}\left(W_\infty\left(\bar{\mathbf{A}}\right)\right)\end{aligned}$$

and

$$\mathbf{\Sigma}^{(q)} = \left[\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}N\right]\times\left[\mathcal{O}\left(q^{2m_2-2}|\lambda_2|^{2q}\right) + \operatorname{Var}\left(W_\infty\left(\bar{\mathbf{A}}\right)\right)\right]\mathbf{1}\mathbf{1}^T.$$

Since $\lim_{q\to\infty}\left[\mathbf{w}_{i,a}^{(q)}\left(\bar{\mathbf{A}}\right)\right]^T = \left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T$ for all $i \in \{1,\ldots,N\}$, by Theorem A.4,

$$\lim_{q\to\infty}\mathbf{\Sigma}^{(q)}\left(\bar{\mathbf{A}},N,n\right) = \left[\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}N\,\mathrm{Var}\,W_\infty\left(\bar{\mathbf{A}}\right)\right]\mathbf{11}^T. \quad \square$$

Proof of Theorem A.6

$\mathrm{Var}\,X\left(\bar{\mathbf{A}},N,n\right) = \frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N}{N-1}N\,\mathrm{Var}\,W\left(\bar{\mathbf{A}}\right)$, so $\mathrm{Var}\,X\left(\bar{\mathbf{A}},N,n\right)$ is maximal when $\mathrm{Var}\,W\left(\bar{\mathbf{A}}\right)$ is maximal. Now,

$$\begin{aligned}\mathrm{Var}\,W\left(\bar{\mathbf{A}}\right) &= \frac{1}{N}\sum_{j=1}^{N}\left(\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j - \frac{1}{N}\right)^2 \\ &= \frac{1}{N}\left[\sum_{j=1}^{N}\left(\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j\right)^2 - \frac{2}{N}\sum_{j=1}^{N}\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j + \sum_{j=1}^{N}\frac{1}{N^2}\right] \\ &= \frac{1}{N}\left[\sum_{j=1}^{N}\left(\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j\right)^2 - \frac{1}{N}\right],\end{aligned}$$

so $\mathrm{Var}\,W\left(\bar{\mathbf{A}}\right)$ is maximal when $\sum_{j=1}^{N}\left(\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j\right)^2$ is maximal. Choose elements for the multiset $\left\{\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j\right\}_{j=1}^{N}$, whose values are constrained so that $\sum_{j=1}^{N}\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j = 1$ and $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j \in [0,1)\ \forall j \in \{1,\ldots,N\}$. Now, for $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_k \leq \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i$, transfer ϵ mass from $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_k$ to $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i$ so that $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_k \longrightarrow \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_k - \epsilon$ and $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i \longrightarrow \left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i + \epsilon$. The constraints continue to be satisfied and $\sum_{j=1}^{N}\left(\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j\right)^2$ increases. $\mathrm{Var}\,W\left(\bar{\mathbf{A}}\right)$ is thus maximal when $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_i = 1$ and $\left[\mathbf{w}\left(\bar{\mathbf{A}}\right)\right]_j = 0,\ \forall j \neq i$. $\square$

Proof of Corollary A.1

Corollary A.1 follows immediately from Theorem A.6 by replacing $\mathbf{w}\left(\bar{\mathbf{A}}\right)$ with $\mathbf{w}_{a,i}\left(\bar{\mathbf{A}}\right)$ and replacing $X\left(\bar{\mathbf{A}},N,n\right)$ with $\widehat{F}_{a,i}\left(\bar{\mathbf{A}},N,n\right)$. $\square$

Proof of Corollary A.2

(1) Start with $[\mathbf{d}_w^-]_j = 1$ and $[\mathbf{d}_w^-]_k = 0$, $\forall k \neq j$. Since $\mathcal{G}(\mathbf{A})$ is undirected, for graph $\mathcal{G}(\bar{\mathbf{A}})$, assign the weight $\frac{1}{N-1}$ to each directed edge from node j to node k, $\forall k \neq j$. Then $[\mathbf{d}_w^-]_k = \frac{1}{N}\left(\frac{1}{N-1}\right)$ $\forall k \neq j$ and $[\mathbf{d}_w^-]_j = \frac{1}{N}(N-1) = 1 - \frac{1}{N}$.

(2) Start with $[\mathbf{d}_w^-]_j = 1$ and $[\mathbf{d}_w^-]_k = 0$, $\forall k \neq j$. Assign an ϵ weight to every self-loop for nodes $k \neq j$ for $\epsilon > 0$ small. Since $\mathcal{G}(\mathbf{A})$ is undirected, introduce the weight $\frac{\epsilon}{N-1}$ for each directed edge in graph $\mathcal{G}(\bar{\mathbf{A}})$ from node j to node k, $\forall k \neq j$. Then $[\mathbf{d}_w^-]_k = \frac{1}{N}\left(\epsilon + \frac{\epsilon}{N-1}\right) = \frac{\epsilon}{N-1}$ and $[\mathbf{d}_w^-]_j = 1 - (N-1)\frac{\epsilon}{N-1} = 1 - \epsilon$.

(3) Part (3) of Corollary A.2 follows immediately from Theorem A.6. Replace the pair $(X(\bar{\mathbf{A}}, N, n), \mathbf{w}(\bar{\mathbf{A}}))$ with $\left(\widehat{F}_{avg}(\bar{\mathbf{A}}, N, n), \mathbf{d}_w^-(\bar{\mathbf{A}})\right)$. Node j has a self-loop to preserve the row-stochasticity of $\bar{\mathbf{A}}$.

(4) Start with $[\mathbf{d}_w^-]_j = 1$ and $[\mathbf{d}_w^-]_k = 0$, $\forall k \neq j$. Assign an ϵ weight to every self-loop for nodes $k \neq j$, for $\epsilon > 0$ small, so that $[\mathbf{d}_w^-]_k = \frac{\epsilon}{N}$, $\forall k \neq j$. Then $[\mathbf{d}_w^-]_j = 1 - (N-1)\left(\frac{\epsilon}{N}\right) = 1 - \frac{N-1}{N}\epsilon$. □

Proof of Corollary A.3

$[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j = \frac{d_j}{\sum_{k=1}^{N} d_k}$. Assuming that every node has a self-loop, $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j$ is maximal when $d_j = (N-1)+1$ and $d_k = 1+1 = 2$ $\forall k \neq j$, so $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j = \frac{N}{N+2(N-1)} = \frac{N}{3N-2}$ and $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_k = \frac{2}{3N-2}$ $\forall k \neq j$. □

Proof of Corollary A.4

$[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j = \frac{d_j^+}{\sum_{k=1}^{N} d_k^+}$. Assuming that every node has a self-loop, since $\bar{\mathbf{A}}$ is strongly connected, $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j$ is maximal when $d_j^+ = N$ and $d_k^+ = 2$ $\forall k \neq j$, so $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_j = \frac{N}{N+(N-1)2} = \frac{N}{3N-2}$ and $[\mathbf{w}_\infty(\bar{\mathbf{A}})]_k = \frac{2}{3N-2}$ $\forall k \neq j$. □

Proof of Theorem A.7

Consider a set of N nodes connected with M undirected edges. If every node has a self-loop, $[\mathbf{w}_\infty]_i = \frac{d_i+1}{\sum_{j=1}^N d_j+1}$, where d_i is the degree of node i arising from non-self-loop edges. If every node lacks a self-loop, $[\mathbf{w}_\infty]_i = \frac{d_i}{\sum_{j=1}^N d_j}$. Given N, n, $\operatorname{Var} \widehat{F}^{(\infty)}(\bar{\mathbf{A}}, N, n)$ is maximal when $\operatorname{Var} W_\infty(\bar{\mathbf{A}}) = \frac{1}{N}\sum_{i=1}^N \left([\mathbf{w}_\infty(\bar{\mathbf{A}})]_i - \frac{1}{N}\right)^2$ is maximal, so we seek a network topology that maximizes this latter quantity. Now, if every node has a self-loop,

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^N \left([\mathbf{w}_\infty(\bar{\mathbf{A}})]_i - \frac{1}{N}\right)^2 &= \frac{1}{N}\sum_{i=1}^N \left(\frac{d_i+1}{\sum_{j=1}^N d_j+1} - \frac{1}{N}\right)^2 \\ &= \frac{1}{N}\left[\frac{\left(\sum_{i=1}^N d_i^2\right)+4M+N}{(2M+N)^2} - \frac{1}{N}\right],\end{aligned}$$

and if every node lacks a self-loop,

$$\frac{1}{N}\sum_{i=1}^N \left([\mathbf{w}_\infty(\bar{\mathbf{A}})]_i - \frac{1}{N}\right)^2 = \frac{1}{N}\sum_{i=1}^N \left(\frac{d_i}{\sum_{j=1}^N d_j} - \frac{1}{N}\right)^2 = \frac{1}{N}\left[\frac{\sum_{i=1}^N d_i^2}{4M^2} - \frac{1}{N}\right].$$

so, given N, M, in both cases we seek a network topology that maximizes $\sum_{i=1}^N d_i^2$.

Lemma A.8 *Given* N, M, $\max_{\{d_i\}_{i=1}^N} \sum_{i=1}^N d_i^2 = \max\{C(N,M), S(N,M)\}$, *where* $C(N,M)$ *is the sum of squared degrees for the simple* $QC(N,M)$ *and* $S(N,M)$ *is the sum of squared degrees for the simple* $QS(N,M)$.

Proof. Lemma A.8 follows from Theorem 2 of Ahlswede and Katona (1978), which characterizes those graphs with N nodes and M edges that have a maximal number of adjacent pairs of edges. Maximizing that quantity reduces to maximizing the graph's sum of squared degrees arising from non-self-loop edges. ■

Theorem A.7 then follows from Lemma A.8. For a particular pair (N, M), if $QC(N,M)$ is inadmissible due to lack of connectedness, $QS(N,M)$ maximizes the sum of squared degrees for the graph (as it satisfies Lemma 2 of Ahlswede and Katona (1978)). $\square$

Proof of Theorem A.8

Suppose that $[\mathbf{w}(\bar{\mathbf{A}}')]^T \neq [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{R}$ for every permutation matrix $\mathbf{R}$. Assign each element $[\mathbf{w}(\bar{\mathbf{A}}')]_j$ an integer $k \in \{1, \ldots, N\}$ so that $[\mathbf{w}(\bar{\mathbf{A}}')]_j^k \leq [\mathbf{w}(\bar{\mathbf{A}}')]_{j'}^{k'}$ whenever $k < k'$, and construct the ordered multiset $\left\{[\mathbf{w}(\bar{\mathbf{A}}')]_j^k\right\}_{k=1}^N$ of weakly increasing weights, in which $[\mathbf{w}(\bar{\mathbf{A}}')]_j^k$ is the k^{th} element of the set. Similarly, assign each element $[\mathbf{w}(\bar{\mathbf{A}})]_j$ an integer $k \in \{1, \ldots, N\}$ so that $[\mathbf{w}(\bar{\mathbf{A}})]_j^k \leq [\mathbf{w}(\bar{\mathbf{A}})]_{j'}^{k'}$ whenever $k < k'$, and construct the ordered multiset $\left\{[\mathbf{w}(\bar{\mathbf{A}})]_j^k\right\}_{k=1}^N$ of weakly increasing weights, in which $[\mathbf{w}(\bar{\mathbf{A}})]_j^k$ is the k^{th} element of the set. Since $[\mathbf{w}(\bar{\mathbf{A}}')]^T \neq [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{R}$ for every possible permutation matrix $\mathbf{R}$, $\left\{[\mathbf{w}(\bar{\mathbf{A}}')]_j^k\right\}_{k=1}^N \neq \left\{[\mathbf{w}(\bar{\mathbf{A}})]_j^k\right\}_{k=1}^N$. Choose the smallest integer $\ell \in \{1, \ldots, N\}$ for which $[\mathbf{w}(\bar{\mathbf{A}}')]_j^\ell \neq [\mathbf{w}(\bar{\mathbf{A}})]_{j'}^\ell$ and $[\mathbf{w}(\bar{\mathbf{A}}')]_j^{\ell-1} = [\mathbf{w}(\bar{\mathbf{A}})]_{j'}^{\ell-1}$. Define $T' = \sum_{k=1}^\ell [\mathbf{w}(\bar{\mathbf{A}}')]_j^k$ and $T = \sum_{k=1}^\ell [\mathbf{w}(\bar{\mathbf{A}})]_j^k$. Set $t \in (\min\{T', T\}, \max\{T', T\})$. Then, when $n = 1$,

$$\left| G_{X(\bar{\mathbf{A}}',N,n)}(t) - G_{X(\bar{\mathbf{A}},N,n)}(t) \right| = \frac{1}{N},$$

so $G_{X(\bar{\mathbf{A}}',N,n)}(t) \neq G_{X(\bar{\mathbf{A}},N,n)}(t)$ for all $t \in \mathbb{R}$ and $n \in \{0, \ldots, N\} \subseteq \mathbb{Z}_+$.

Now suppose that $[\mathbf{w}(\bar{\mathbf{A}}')]^T = [\mathbf{w}(\bar{\mathbf{A}})]^T \mathbf{R}$ for some permutation matrix $\mathbf{R}$.

Then,

$$\begin{aligned}
G_{X(\bar{\mathbf{A}}',N,n)}(t) &= \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{x(\bar{\mathbf{A}}',\mathbf{b},N,n)\le t} \\
&= \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{[\mathbf{w}(\bar{\mathbf{A}}')]^T\mathbf{b}(N,n)\le t} \\
&= \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{[\mathbf{w}(\bar{\mathbf{A}})]^T\mathbf{R}\mathbf{b}(N,n)\le t} \\
&= \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{[\mathbf{w}(\bar{\mathbf{A}})]^T\mathbf{b}(N,n)\le t} \\
&= \frac{1}{|\mathcal{B}(N,n)|}\sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{x(\bar{\mathbf{A}},\mathbf{b},N,n)\le t} \\
&= G_{X(\bar{\mathbf{A}},N,n)}(t)
\end{aligned}$$

for all $t \in \mathbb{R}$ and $n \in \{0,\ldots,N\} \subseteq \mathbb{Z}_+$. $\square$

Proof of Corollary A.5

Replace the pair $\left(\mathbf{w}(\bar{\mathbf{A}}), G_{X(\bar{\mathbf{A}},N,n)}(t)\right)$ in the proof of Theorem A.8 with each pair specified in the statement of Corollary A.5. For example, to demonstrate that $G_{\widehat{F}_{avg}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$, replace $\left(\mathbf{w}(\bar{\mathbf{A}}), G_{X(\bar{\mathbf{A}},N,n)}(t)\right)$ in the proof of Theorem A.8 with $\left(\mathbf{d}_w^-(\bar{\mathbf{A}}), G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)\right)$. Corollary A.5 then follows. $\square$

Proof of Theorem A.9

To demonstrate (1): From Corollary A.5, $G_{\widehat{F}_i(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_i(\bar{\mathbf{A}},N,n)}(t)$ if and only if $[\mathbf{w}_{a,i}(\bar{\mathbf{A}}')]^T = [\mathbf{w}_{a,i}(\bar{\mathbf{A}})]^T\mathbf{R}$ for some permutation matrix $\mathbf{R}$. Since $[\mathbf{w}_{a,i}(\bar{\mathbf{A}})]^T \equiv [\bar{\mathbf{A}}]_{i*}$, $[\mathbf{w}_{a,i}(\bar{\mathbf{A}}')]^T = [\mathbf{w}_{a,i}(\bar{\mathbf{A}})]^T\mathbf{R}$ if and only if $[\bar{\mathbf{A}}']_{i*} = [\bar{\mathbf{A}}]_{i*}\mathbf{R}$. Similarly, since $\left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]^T \equiv [\bar{\mathbf{A}}^q]_{i*}$, $G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_i^{(q)}(\bar{\mathbf{A}},N,n)}(t)$ if and only if $\left[(\bar{\mathbf{A}}')^q\right]_{i*} = [\bar{\mathbf{A}}^q]_{i*}\mathbf{R}$.

To demonstrate (2): From Corollary A.5, $G_{\widehat{F}_{avg}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ if and only if $[\mathbf{d}_{\overline{w}}^{-}(\bar{\mathbf{A}}')]^T = [\mathbf{d}_{\overline{w}}^{-}(\bar{\mathbf{A}})]^T \mathbf{R}$ for some permutation matrix $\mathbf{R}$, that is, if and only if $\mathbf{1}^T\bar{\mathbf{A}}' = \mathbf{1}^T\bar{\mathbf{A}}\mathbf{R}$. Similarly, $G_{\widehat{F}^{(q)}_{avg}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}^{(q)}_{avg}(\bar{\mathbf{A}},N,n)}(t)$ if and only if $\left[\mathbf{d}_{\overline{w}}^{-(q)}(\bar{\mathbf{A}}')\right]^T = \left[\mathbf{d}_{\overline{w}}^{-(q)}(\bar{\mathbf{A}})\right]^T \mathbf{R}$ for some permutation matrix $\mathbf{R}$, that is, if and only if $\mathbf{1}^T(\bar{\mathbf{A}}')^q = \mathbf{1}^T\bar{\mathbf{A}}^q\mathbf{R}$.

To demonstrate (3): From Corollary A.5, $G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}}',N,n)}(t) = G_{\widehat{F}^{(\infty)}(\bar{\mathbf{A}},N,n)}(t)$ if and only if $[\mathbf{w}_\infty(\bar{\mathbf{A}}')]^T = [\mathbf{w}_\infty(\bar{\mathbf{A}})]^T \mathbf{R}$ for some permutation matrix $\mathbf{R}$. $[\mathbf{w}_\infty(\bar{\mathbf{A}})]^T = [\mathbf{w}_\infty(\bar{\mathbf{A}})]^T \bar{\mathbf{A}}$, so $[\mathbf{w}_\infty(\bar{\mathbf{A}}')]^T = [\mathbf{w}_\infty(\bar{\mathbf{A}})]^T \mathbf{R}$ if and only if $[\mathbf{w}_\infty(\bar{\mathbf{A}}')]^T \mathbf{R}^T = [\mathbf{w}_\infty(\bar{\mathbf{A}}')]^T \mathbf{R}^T\bar{\mathbf{A}}$, or equivalently, $[\mathbf{w}_\infty(\bar{\mathbf{A}}')]^T = [\mathbf{w}_\infty(\bar{\mathbf{A}}')]^T \mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}$, that is, if and only if $\bar{\mathbf{A}}'$ and $\mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}$ have a common dominant left eigenpair. Note that both $\bar{\mathbf{A}}$ and $\bar{\mathbf{A}}'$ are row-stochastic and primitive. Since $\bar{\mathbf{A}}$ is row-stochastic and primitive, $\mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}$ is also row-stochastic and primitive, with primitivity of $\mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}$ observable from the structural isomorphism of $\mathcal{G}(\bar{\mathbf{A}})$ and $\mathcal{G}\left(\mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}\right)$. Therefore, the dominant left eigenpair of $\mathbf{R}^T\bar{\mathbf{A}}\mathbf{R}$ is also unique, with the unit-normalized left eigenvector paired to an eigenvalue of value 1. □

Proof of Theorem A.10

Demonstrating that $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ if and only if there exists a permutation matrix $\mathbf{R}$ such that $\bar{\mathbf{A}}' = \bar{\mathbf{A}}\mathbf{R}$: Suppose that $\bar{\mathbf{A}}' \neq \bar{\mathbf{A}}\mathbf{R}$ for every permutation matrix $\mathbf{R}$. Then multiset $\{[\bar{\mathbf{A}}']_{*1}, \ldots, [\bar{\mathbf{A}}']_{*N}\} \neq \{[\bar{\mathbf{A}}]_{*1}, \ldots, [\bar{\mathbf{A}}]_{*N}\}$. Set $n = 1$. $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_{[\bar{\mathbf{A}}']_{*j}\leq\mathbf{t}}$ and $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t}) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_{[\bar{\mathbf{A}}]_{*j}\leq\mathbf{t}}$, so $G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) \neq G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$.

Now suppose that $\bar{\mathbf{A}}' = \bar{\mathbf{A}}\mathbf{R}$ for some permutation matrix $\mathbf{R}$. Then

$$\begin{aligned}
G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) &= \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\bar{\mathbf{A}}'\mathbf{b}(N,n)\le\mathbf{t}} \\
&= \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\bar{\mathbf{A}}\mathbf{R}\mathbf{b}(N,n)\le\mathbf{t}} \\
&= \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\bar{\mathbf{A}}\mathbf{b}(N,n)\le\mathbf{t}} \\
&= G_{\widehat{\mathbf{F}}(\bar{\mathbf{A}},N,n)}(\mathbf{t}).
\end{aligned}$$

To demonstrate that $G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}}',N,n)}(\mathbf{t}) = G_{\widehat{\mathbf{F}}^{(q)}(\bar{\mathbf{A}},N,n)}(\mathbf{t})$ if and only if there exists a permutation matrix $\mathbf{R}$ such that $(\bar{\mathbf{A}}')^q = \bar{\mathbf{A}}^q\mathbf{R}$, replace $\bar{\mathbf{A}}'$ with $(\bar{\mathbf{A}}')^q$ and replace $\bar{\mathbf{A}}$ with $(\bar{\mathbf{A}})^q$ in the above proof. □

Proof of Theorem A.11

Note that

$$\begin{aligned}
G_{\widehat{\mathbf{F}}_S(\bar{\mathbf{A}},N,n)}(\mathbf{t}) &= \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\mathbf{S}\,\hat{\mathbf{f}}(\bar{\mathbf{A}},\mathbf{b},N,n)\le\mathbf{t}} \\
&= \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\mathbf{S}\,\bar{\mathbf{A}}\mathbf{b}(N,n)\le\mathbf{t}} \\
&= G_{\widehat{\mathbf{F}}(\mathbf{S}\bar{\mathbf{A}},N,n)}(\mathbf{t}).
\end{aligned}$$

Therefore, replace $\bar{\mathbf{A}}$ with $\mathbf{S}\bar{\mathbf{A}}$ and replace $\bar{\mathbf{A}}^q$ with $\mathbf{S}\bar{\mathbf{A}}^q$ in the proof of Theorem A.10. Theorem A.11 then follows. □

Proof of Theorem A.12

The expression that approximates $\mathbf{w}_\infty^T(\epsilon)$ following a perturbation $\bar{\mathbf{A}}(\epsilon)$ applies the work of Conlisk (1985) on perturbation theory for finite Markov chains. Define the fundamental matrix (Kemeny and Snell, 1960) $\mathbf{Z} = \left(\mathbf{I} - \bar{\mathbf{A}} + \mathbf{1}\mathbf{p}^T\right)^{-1}$, for any $N \times 1$ vector $\mathbf{p}$ such that $\mathbf{p}^T\mathbf{1} \neq 0$. Matrix $\mathbf{Z}$ always exists because $\det\left(\mathbf{I} - \bar{\mathbf{A}} + \mathbf{1}\mathbf{p}^T\right) \neq 0$ (see Conlisk, 1985, and Theorem 29 in Brauer, 1952). For the perturbation $\bar{\mathbf{A}}(\epsilon) = \bar{\mathbf{A}} + \epsilon\,\mathbf{E}$,

$$\left.\frac{\partial \mathbf{Z}(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = \mathbf{Z}\left(\left.\frac{\partial \bar{\mathbf{A}}(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0}\right)\mathbf{Z} = \mathbf{ZEZ}.$$

Since $\mathbf{w}_\infty^T = \mathbf{p}^T\mathbf{Z}$,

$$\left.\frac{\partial \mathbf{w}_\infty^T(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = \mathbf{b}^T\left(\left.\frac{\partial \mathbf{Z}(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0}\right) = \mathbf{b}^T\mathbf{ZEZ} = \mathbf{w}_\infty^T\mathbf{EZ}.$$

Next, to obtain the exact expression for $\left[\mathbf{d}_w^-(\epsilon)\right]^T$, observe that

$$\left[\mathbf{d}_w^-(\epsilon)\right]^T = \frac{1}{N}\mathbf{1}^T\bar{\mathbf{A}}(\epsilon) = \frac{1}{N}\mathbf{1}^T\left(\bar{\mathbf{A}} + \epsilon\,\mathbf{E}\right) = \left[\mathbf{d}_w^-\right]^T + \frac{\epsilon}{N}\left(\mathbf{1}^T\mathbf{E}\right),$$

and to obtain the exact expression for $\left[\mathbf{w}_{a,i}(\epsilon)\right]^T$, $\forall i \in \{1,\ldots,N\}$, observe that

$$\left[\mathbf{w}_{a,i}(\epsilon)\right]^T = \left[\bar{\mathbf{A}}(\epsilon)\right]_{i*} = \left[\bar{\mathbf{A}} + \epsilon\,\mathbf{E}\right]_{i*} = \left[\mathbf{w}_{a,i}\right]^T + \epsilon\,\left[\mathbf{E}\right]_{i*}.$$

Finally, to obtain the exact expression for $\left[\mathbf{d}_w^{-(q)}(\epsilon)\right]^T$, observe that

$$\left[\mathbf{d}_w^{-(q)}(\epsilon)\right]^T = \frac{1}{N}\mathbf{1}^T\bar{\mathbf{A}}^q(\epsilon) = \frac{1}{N}\mathbf{1}^T\left(\bar{\mathbf{A}}^q + \epsilon\,\mathbf{E}\right) = \left[\mathbf{d}_w^{-(q)}\right]^T + \frac{\epsilon}{N}\left(\mathbf{1}^T\mathbf{E}\right),$$

and to obtain the exact expression for $\left[\mathbf{w}_{a,i}^{(q)}(\epsilon)\right]^T$, $\forall i \in \{1,\ldots,N\}$, observe that

$$\left[\mathbf{w}_{a,i}^{(q)}(\epsilon)\right]^T = \left[\bar{\mathbf{A}}^q(\epsilon)\right]_{i*} = \left[\bar{\mathbf{A}}^q + \epsilon\,\mathbf{E}\right]_{i*} = \left[\mathbf{w}_{a,i}^{(q)}\right]^T + \epsilon\,\left[\mathbf{E}\right]_{i*}. \quad \square$$

Proof of Corollary A.6

For part (1), set $\mathbf{E} = \mathbf{e}_i\left(\mathbf{e}_j^T - \mathbf{e}_k^T\right)$, where $\mathbf{e}_i$ is the $N \times 1$ unit vector whose i^{th} element equals 1 and all other elements equal zero. From Theorem A.12, the r^{th}

element of $\left.\frac{\partial \mathbf{w}_\infty^T(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0}$ is

$$\left.\frac{\partial\left[\mathbf{w}_\infty^T(\epsilon)\right]_r}{\partial \epsilon}\right|_{\epsilon=0} = \mathbf{w}_\infty^T \mathbf{E}\mathbf{Z}\mathbf{e}_r = \left[\mathbf{w}_\infty^T\right]_i \left([\mathbf{Z}]_{jr} - [\mathbf{Z}]_{kr}\right).$$

Define the $N \times N$ mean first passage matrix $\mathbf{M} = \mathbf{1}\mathbf{1}^T + \bar{\mathbf{A}}\left(\mathbf{M} - \mathbf{M}_{diag}\right)$, where $\mathbf{M}_{diag} = \text{diag}\left([\mathbf{M}]_{11}, \ldots, [\mathbf{M}]_{NN}\right)$. From Conlisk (1985), for the vector $\mathbf{w}_\infty^T$,

$$[\mathbf{M}]_{rs} = \begin{cases} \frac{1}{[\mathbf{w}_\infty^T]_s} & \text{for } r = s, \text{ and} \\ \frac{[\mathbf{Z}]_{ss} - [\mathbf{Z}]_{rs}}{[\mathbf{w}_\infty^T]_s} & \text{for } r \neq s. \end{cases}$$

Since the elements of $\mathbf{M}$ and $\mathbf{w}_\infty^T$ are strictly positive, $[\mathbf{Z}]_{ss} > [\mathbf{Z}]_{rs}$ for $r \neq s$. Thus, setting $r = j$,

$$\left.\frac{\partial\left[\mathbf{w}_\infty^T(\epsilon)\right]_j}{\partial \epsilon}\right|_{\epsilon=0} = \left[\mathbf{w}_\infty^T\right]_i \left([\mathbf{Z}]_{jj} - [\mathbf{Z}]_{kj}\right) > 0,$$

and setting $r = k$,

$$\left.\frac{\partial\left[\mathbf{w}_\infty^T(\epsilon)\right]_k}{\partial \epsilon}\right|_{\epsilon=0} = \left[\mathbf{w}_\infty^T\right]_i \left([\mathbf{Z}]_{jk} - [\mathbf{Z}]_{kk}\right) < 0.$$

For part (2), first set

$$\mathbf{E} = \sum_{\substack{k=1 \\ k \neq j}} \alpha_k \mathbf{e}_i \left(\mathbf{e}_j^T - \mathbf{e}_k^T\right),$$

with $\alpha_k \geq 0$ for every $k \in \{1, \ldots, N\}$, $k \neq j$, so that $[\mathbf{E}]_{ij} > 0$, $[\mathbf{E}]_{ik} \leq 0$ for all $k \neq j$, and $\sum_{k=1}^{N}[\mathbf{E}]_{ik}=0$. From Theorem A.12, the r^{th} element of $\left.\frac{\partial \mathbf{w}_{\infty}^{T}(\epsilon)}{\partial \epsilon}\right|_{\epsilon=0}$ is

$$\left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{r}}{\partial \epsilon}\right|_{\epsilon=0}=\mathbf{w}_{\infty}^{T}\left[\sum_{\substack{k=1\\k\neq j}}^{N}\alpha_{k}\mathbf{e}_{i}\left(\mathbf{e}_{j}^{T}-\mathbf{e}_{k}^{T}\right)\right]\mathbf{Z}\mathbf{e}_{r}=\left[\sum_{\substack{k=1\\k\neq j}}^{N}\alpha_{k}\mathbf{w}_{\infty}^{T}\mathbf{e}_{i}\left(\mathbf{e}_{j}^{T}-\mathbf{e}_{k}^{T}\right)\right]\mathbf{Z}\mathbf{e}_{r}$$
$$=\sum_{\substack{k=1\\k\neq j}}^{N}\alpha_{k}\left[\mathbf{w}_{\infty}^{T}\right]_{i}\left([\mathbf{Z}]_{jr}-[\mathbf{Z}]_{kr}\right).$$

Setting $r=j$,

$$\left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{j}}{\partial \epsilon}\right|_{\epsilon=0}=\sum_{\substack{k=1\\k\neq j}}^{N}\alpha_{k}\left[\mathbf{w}_{\infty}^{T}\right]_{i}\left([\mathbf{Z}]_{jj}-[\mathbf{Z}]_{kj}\right)>0.$$

Next consider

$$\mathbf{E}=-\sum_{\substack{k=1\\k\neq j}}\alpha_{k}\mathbf{e}_{i}\left(\mathbf{e}_{j}^{T}-\mathbf{e}_{k}^{T}\right),$$

with $\alpha_k \geq 0$ for every $k\in\{1,\ldots,N\}$, $k\neq j$, so that $[\mathbf{E}]_{ij}<0$, $[\mathbf{E}]_{ik}\geq 0$ for all $k\neq j$, and $\sum_{k=1}^{N}[\mathbf{E}]_{ik}=0$. Then,

$$\left.\frac{\partial\left[\mathbf{w}_{\infty}^{T}(\epsilon)\right]_{j}}{\partial \epsilon}\right|_{\epsilon=0}=-\sum_{\substack{k=1\\k\neq j}}^{N}\alpha_{k}\left[\mathbf{w}_{\infty}^{T}\right]_{i}\left([\mathbf{Z}]_{jj}-[\mathbf{Z}]_{kj}\right)<0.\ \square$$

Proof of Theorem A.13

The variance of $X(\bar{\mathbf{A}},N,n)$ follows directly from Theorem 1.9, and the asymptotic expansion for $X(\bar{\mathbf{A}},N,n)$ follows directly from Theorem 1.13. With $B_i\sim\text{Bern}\left(\frac{n}{N}\right)$,

$$\begin{aligned}EX\left(\bar{\mathbf{A}}(\epsilon),N,n\right)&=E\left[w_1(\epsilon)B_1+w_2(\epsilon)B_2+\cdots+w_N(\epsilon)B_M\right]\\&=\frac{n}{N}\sum_{i=1}^{N}w_i(\epsilon)\\&=\frac{n}{N}.\ \square\end{aligned}$$

Appendix B

Appendix to Chapter 2

B.1 Proofs

Proof of Lemma 2.1

First method of proof: $E\widehat{F}_{avg}(\mathbf{Z}, N, n) = E\left(\left[\mathbf{d}_w^-(\mathbf{Z})\right]^T \mathbf{B}(N, n)\right)$, where $\mathbf{B}(N, n)$ is a random vector whose elements are $B_i \sim \text{Bern}\left(\frac{n}{N}\right), i \in \{1, \ldots, N\}$ and $\left[\mathbf{d}_w^-(\mathbf{Z})\right]^T \mathbf{1} = k$. Therefore,

$$\begin{aligned} E\widehat{F}_{avg}(\mathbf{Z}, N, n) &= E\left(\left[\mathbf{d}_w^-(\mathbf{Z})\right]_1 B_1 + \left[\mathbf{d}_w^-(\mathbf{Z})\right]_2 B_2 + \ldots + \left[\mathbf{d}_w^-(\mathbf{Z})\right]_N B_N\right) \\ &= \sum_{i=1}^{N} \left[\mathbf{d}_w^-(\mathbf{Z})\right]_i EB_i \\ &= \frac{kn}{N}. \quad \square \end{aligned}$$

Second method of proof:

$$\begin{aligned}
E\widehat{F}_{avg}(\mathbf{Z}, N, n) &= \frac{1}{|\mathcal{B}(N,n)|} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \left[\mathbf{d}_w^-(\mathbf{Z})\right]^T \mathbf{b}(N,n) \\
&= \frac{1}{|\mathcal{B}(N,n)|} \sum_{i=1}^{N} |\{\mathbf{b} \in \mathcal{B}(N,n) : b_i = 1\}| \left[\mathbf{d}_w^-(\mathbf{Z})\right]_i \\
&= \frac{|\{\mathbf{b} \in \mathcal{B}(N,n) : b_i = 1\}|}{|\mathcal{B}(N,n)|} \sum_{i=1}^{N} \left[\mathbf{d}_w^-(\mathbf{Z})\right]_i = \left(\frac{\binom{N}{n} \times \frac{n}{N}}{\binom{N}{n}}\right) \times k = \frac{kn}{N}.
\end{aligned}$$

By symmetry, $|\{\mathbf{b} \in \mathcal{B}(N,n) : b_i = 1\}| = \left|\{\mathbf{b} \in \mathcal{B}(N,n) : b_j = 1\}\right|$, $\forall i, j \in \{1, \ldots, N\}$. □

Proof of Lemma 2.2

The proof of this Lemma immediately follows from the proof of Theorem 1.9. In Chapter 1 of this dissertation, $\text{Var}\, D_w^-(\mathbf{Z}) = \frac{1}{N}\sum_{i=1}^{N}\left([\mathbf{d}_w^-(\mathbf{Z})]_i - \frac{1}{N}\right)^2$ because $\mathbf{1}^T\mathbf{d}_w^-(\mathbf{Z}) = 1$, while in the present chapter, $\text{Var}\, D_w^-(\mathbf{Z}) = \frac{1}{N}\sum_{i=1}^{N}\left([\mathbf{d}_w^-(\mathbf{Z})]_i - \frac{k}{N}\right)^2$ because $\mathbf{1}^T\mathbf{d}_w^-(\mathbf{Z}) = k$. □

Proof of Lemma 2.3

The proof of this Lemma immediately follows from the proof of Theorem 1.10. □

Proof of Lemma 2.4

The scalar $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n) = [\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{b}(N,n)$ is invariant to configuration if and only if $[\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{b}(N,n) = [\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{b}'(N,n)$ for all $\mathbf{b}(N,n), \mathbf{b}'(N,n) \in \mathcal{B}(N,n)$, with this relation holding for each integer $n \in [0, N]$. Let $n = 1$, and define $\mathbf{e}_i$ to be the i^{th} unit vector whose i^{th} element equals 1 and all other elements equal zero. Then $[\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{b}(N,1) = [\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{b}'(N,1)$ if and only if $[\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{e}_i =$

$[\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{e}_j$, that is, if and only if $[\mathbf{d}_w^-(\mathbf{Z})]_i = [\mathbf{d}_w^-(\mathbf{Z})]_j$ for all $i, j \in \{1, \ldots, N\}$. Since $\mathbf{1}^T \mathbf{d}_w^-(\mathbf{Z}) = k$, $[\mathbf{d}_w^-(\mathbf{Z})]_i = \frac{k}{N}$ for all $i \in \{1, \ldots, N\}$. Given that $\mathbf{d}_w^-(\mathbf{Z}) = \frac{k}{N}\mathbf{1}$ when $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n)$ is invariant to configuration, $\widehat{f}_{avg}(\mathbf{Z}, \mathbf{b}, N, n) = [\mathbf{d}_w^-(\mathbf{Z})]^T \mathbf{b}(N, n) = \frac{k}{N}\mathbf{1}^T \mathbf{b}(N, n) = \frac{kn}{N}$. □

Proof of Lemma 2.5

The proof of this Lemma immediately follows from the proof of Theorem 1.13. □

Proof of Lemma 2.6

The proof of Lemma 2.6 makes use of the following result from Erdös and Rényi (1959), with notation modified for the present work:

Lemma B.1 (Erdös and Rényi (1959), Theorem 1) *Consider the infinite triangular matrix of real elements*

$$\begin{array}{lllllll} w'_{11} & & & & & & \\ w'_{21} & w'_{22} & & & & & \\ \cdot & \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \\ w'_{N1} & w'_{N2} & \cdot & \cdot & \cdot & w'_{NN} & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

with $\mathbf{w}'_N$ denoting the N^{th} row of the matrix and $\sum_{j=1}^{N} w'_{Nj} = 0$. For any real value t, determine $T\left(\mathbf{w}'_N, N, n, t\right)$, that is, the total number of sums

$$a\left(\mathbf{w}'_N, N, n\right) = w'_{Ni_1} + w'_{Ni_2} + \cdots + w'_{Ni_n}, \qquad 1 \leq i_1 < i_2 < \cdots < i_n \leq N,$$

whose value does not exceed $t\sigma\left(\mathbf{w}'_N, N, n\right) \equiv t\sqrt{\frac{n}{N}\left(1 - \frac{n}{N}\right)\sum_{j=1}^{N} w'^2_{Nj}}$. Let CDF

$G_{\frac{A\left(\mathbf{w}_N',N,n\right)}{\sigma\left(\mathbf{w}_N',N,n\right)}}(t) = \frac{T\left(\mathbf{w}_N',N,n,t\right)}{\binom{N}{n}}$. *With*

$$\kappa\left(\mathbf{w}_N', N, n, \epsilon'\right) \equiv \frac{1}{\sum_{j=1}^{N} w_{Nj}'^2} \sum_{\substack{j\in\{1,\ldots,N\}\ s.t.\\ \left|w_{Nj}'\right|>\epsilon'\sigma\left(\mathbf{w}_N',N,n\right)}} w_{Nj}'^2$$

if $\lim_{N\to\infty} \kappa\left(\mathbf{w}_N', N, n, \epsilon'\right) = 0$ *for any* $\epsilon' > 0$, *then* $\lim_{N\to\infty} G_{\frac{A}{\sigma}}(t) = \Phi(t)$ *for any real* t, *where* $\Phi(\cdot)$ *denotes the standard normal CDF.*

For a given population size N, set $\mathbf{w}_N' = \mathbf{d}_{N,w}^{-}(\mathbf{Z}) - \frac{k}{N}$, where $\mathbf{d}_{N,w}^{-}(\mathbf{Z})$ is the vector of average weighted in-degrees ($\mathbf{d}_w^{-}(\mathbf{Z})$) discussed in the text, and subscript N is added to make the population size explicit. Then $\sum_{j=1}^{N} w_{Nj}' = \sum_{j=1}^{N}\left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_j - \frac{k}{N}\right) = 0$. Scalar quantity

$$\begin{aligned}\widehat{f}_{avg}(\mathbf{Z},\mathbf{b},N,n) - \frac{kn}{N} &= \left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]^T \mathbf{b}(N,n) - \frac{kn}{N} \\ &= \left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_{i_1} - \frac{k}{N}\right) + \left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_{i_2} - \frac{k}{N}\right) \\ &\qquad + \cdots + \left(\left[\mathbf{d}_{N,w}^{-}(\mathbf{Z})\right]_{i_n} - \frac{k}{N}\right) \\ &= w_{Ni_1}' + w_{Ni_2}' + \cdots + w_{Ni_n}',\end{aligned}$$

where $1 \le i_1 < i_2 < \cdots < i_n \le N$, given a configuration $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$. Thus, by Lemma B.1,

$$
\begin{aligned}
G_{\frac{A\left(\mathbf{w}'_N,N,n\right)}{\sigma\left(\mathbf{w}'_N,N,n\right)}}(t) &= \frac{T\left(\mathbf{w}'_N,N,n,t\right)}{\binom{N}{n}} \\
&= \frac{1}{\binom{N}{n}} \sum_{\forall 1\le i_1<i_2<\cdots<i_n\le N} \mathbb{1}_{w'_{Ni_1}+w'_{Ni_2}+\cdots+w'_{Ni_n}\le t\sigma\left(\mathbf{w}'_N,N,n\right)} \\
&= \frac{1}{\binom{N}{n}} \sum_{\mathbf{b}(N,n)\in\mathcal{B}(N,n)} \mathbb{1}_{\widehat{f}_{avg}(\mathbf{Z},\mathbf{b},N,n)-\frac{kn}{N}\le t\sigma\left(\mathbf{w}'_N,N,n\right)} \\
&= G_{\frac{\widehat{f}_{avg}(\mathbf{Z},N,n)-\frac{kn}{N}}{\sigma\left(\mathbf{w}'_N,N,n\right)}}(t),
\end{aligned}
$$

so $\lim_{N\to\infty} G_{\frac{\widehat{f}_{avg}(\mathbf{Z},N,n)-\frac{kn}{N}}{\sigma\left(\mathbf{w}'_N,N,n\right)}}(t) = \Phi(t)$, where $\sigma\left(\mathbf{w}'_N,N,n\right) = \sqrt{\frac{n}{N}\left(1-\frac{n}{N}\right)\sum_{j=1}^{N}\left(\left[\mathbf{d}^{-}_{N,w}(\mathbf{Z})\right]_j-\frac{k}{N}\right)^2}$. □

Proof of Proposition 2.1

This Proposition follows immediately from Equations 2.2-2.5 and Lemma 2.1. □

Proof of Proposition 2.2

This Proposition follows immediately from Equations 2.2-2.5 and Lemma 2.2. □

Proof of Proposition 2.3

This Proposition follows immediately from Equations 2.2-2.5 and Lemma 2.3. □

Proof of Proposition 2.4

Given Equations 2.2-2.5, $y_{agg}(\mathbf{Z},\mathbf{b},N,n,0)$, $m(\mathbf{Z},\mathbf{b},N,n,0)$, $y_{agg}(\mathbf{Z},\mathbf{b},N,n,1)$, and $m(\mathbf{Z},\mathbf{b},N,n,1)$ are all invariant to configuration if and only if $\widehat{f}_{avg}(\mathbf{Z},\mathbf{b},N,n)$ is invariant to configuration. This Proposition then follows from Lemma 2.4. □

Proof of Proposition 2.5

Given Equation 2.2 for $Y_{agg}(\mathbf{Z}, N, n, 0)$:

$$\begin{aligned} G_{\frac{Y_{agg}(\mathbf{Z},N,n,0)-EY_{agg}(\mathbf{Z},N,n,0)}{\left(\mathrm{Var}\,Y_{agg}(\mathbf{Z},N,n,0)\right)^{1/2}}}(t) &= \Pr\left[\frac{Y_{agg}(\mathbf{Z},N,n,0)-EY_{agg}(\mathbf{Z},N,n,0)}{\left(\mathrm{Var}\,Y_{agg}(\mathbf{Z},N,n,0)\right)^{1/2}} \le t\right] \\ &= \Pr\left[\frac{y_{agg}^{no}+\gamma_1\frac{N^2\epsilon}{N-n}\left(\widehat{F}_{avg}(\mathbf{Z},N,n)-\frac{kn}{N}\right)-y_{agg}^{no}}{\gamma_1\frac{N^2\epsilon}{N-n}\left(\mathrm{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}} \le t\right] \\ &= G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\mathrm{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}}(t)\,. \end{aligned}$$

Given Equation 2.3 for $M(\mathbf{Z}, N, n, 0)$:

$$\begin{aligned} G_{\frac{M(\mathbf{Z},N,n,0)-EM(\mathbf{Z},N,n,0)}{\left(\mathrm{Var}\,M(\mathbf{Z},N,n,0)\right)^{1/2}}}(t) &= \Pr\left[\frac{M(\mathbf{Z},N,n,0)-EM(\mathbf{Z},N,n,0)}{\left(\mathrm{Var}\,M(\mathbf{Z},N,n,0)\right)^{1/2}} \le t\right] \\ &= \Pr\left[\frac{\gamma_1\frac{N^2}{N-n}\left(\widehat{F}_{avg}(\mathbf{Z},N,n)-\frac{kn}{N}\right)}{\gamma_1\frac{N^2}{N-n}\left(\mathrm{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}} \le t\right] \\ &= G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\mathrm{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}}(t)\,. \end{aligned}$$

Given Equation 2.4 for $Y_{agg}(\mathbf{Z}, N, n, 1)$:

$$\begin{aligned} G_{\frac{Y_{agg}(\mathbf{Z},N,n,1)-EY_{agg}(\mathbf{Z},N,n,1)}{\left(\mathrm{Var}\,Y_{agg}(\mathbf{Z},N,n,1)\right)^{1/2}}}(t) &= \Pr\left[\frac{Y_{agg}(\mathbf{Z},N,n,1)-EY_{agg}(\mathbf{Z},N,n,1)}{\left(\mathrm{Var}\,Y_{agg}(\mathbf{Z},N,n,1)\right)^{1/2}} \le t\right] \\ &= \Pr\left[\frac{y_{agg}^{no}+\gamma_1 N\epsilon\widehat{F}_{avg}(\mathbf{Z},N,n)-y_{agg}^{no}-\gamma_1 kn\epsilon}{\gamma_1 N\epsilon\left(\mathrm{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}} \le t\right] \\ &= G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\mathrm{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}}(t)\,. \end{aligned}$$

Given Equation 2.5 for $M(\mathbf{Z},N,n,1)$:

$$G_{\frac{M(\mathbf{Z},N,n,1)-EM(\mathbf{Z},N,n,1)}{(\operatorname{Var} M(\mathbf{Z},N,n,1))^{1/2}}}(t) = \Pr\left[\frac{M(\mathbf{Z},N,n,1)-EM(\mathbf{Z},N,n,1)}{(\operatorname{Var} M(\mathbf{Z},N,n,1))^{1/2}} \le t\right]$$
$$= \Pr\left[\frac{\gamma_1 N\widehat{F}_{avg}(\mathbf{Z},N,n)-\gamma_1 kn}{\gamma_1 N\left(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}} \le t\right]$$
$$= G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n))^{1/2}}}(t).$$

The Proposition then follows from Lemma 2.5. □

Proof of Proposition 2.6

From the proof of Proposition 2.5, we have that

$$G_{\frac{Y_{agg}(\mathbf{Z},N,n,\ell)-EY_{agg}(\mathbf{Z},N,n,\ell)}{(\operatorname{Var} Y_{agg}(\mathbf{Z},N,n,\ell))^{1/2}}}(t) = G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n))^{1/2}}}(t) \text{ and}$$

$$G_{\frac{M(\mathbf{Z},N,n,\ell)-EM(\mathbf{Z},N,n,\ell)}{(\operatorname{Var} M(\mathbf{Z},N,n,\ell))^{1/2}}}(t) = G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{(\operatorname{Var}\widehat{F}_{avg}(\mathbf{Z},N,n))^{1/2}}}(t)$$

for $\ell \in \{0,1\}$. Therefore, by Lemma 2.6, if $\lim_{N\to\infty}\kappa_N(\epsilon') = 0$ for any $\epsilon' > 0$, then

$$\lim_{N\to\infty} G_{\frac{Y_{agg}(\mathbf{Z},N,n,\ell)-EY_{agg}(\mathbf{Z},N,n,\ell)}{(\operatorname{Var} Y_{agg}(\mathbf{Z},N,n,\ell))^{1/2}}}(t) = \Phi(t) \text{ and}$$

$$\lim_{N\to\infty} G_{\frac{M(\mathbf{Z},N,n,\ell)-EM(\mathbf{Z},N,n,\ell)}{(\operatorname{Var} M(\mathbf{Z},N,n,\ell))^{1/2}}}(t) = \Phi(t)$$

for $\ell \in \{0,1\}$ and for all real t. □

Proof of Proposition 2.7

The statements that $\Pr\left[Y_{agg}(\mathbf{Z},N,n,0) < y^{no}_{agg}\right] = \Pr\left[M(\mathbf{Z},N,n,0) < 0\right] = \Pr\left[\widehat{F}_{avg}(\mathbf{Z},N,n) < \frac{kn}{N}\right]$ and $\Pr\left[Y_{agg}(\mathbf{Z},N,n,1) < y^{no}_{agg}\right] = \Pr\left[M(\mathbf{Z},N,n,1) < 0\right] = \Pr\left[\widehat{F}_{avg}(\mathbf{Z},N,n) < 0\right]$ follow from Equations 2.2-2.5. From Lemma 2.5, provided that condition (c) holds, $G_{\frac{\widehat{F}_{avg}(\mathbf{Z},N,n)-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\text{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}}(t) \approx J(\mathbf{Z},N,n,t)$ and therefore $G_{\widehat{F}_{avg}(\mathbf{Z},N,n)}(t) \approx J\left(\mathbf{Z},N,n,\frac{t-E\widehat{F}_{avg}(\mathbf{Z},N,n)}{\left(\text{Var}\,\widehat{F}_{avg}(\mathbf{Z},N,n)\right)^{1/2}}\right)$. □

Proof of Proposition 2.8

For every $n \in \{1,\ldots,N-1\}$, we show that there is a positive probability of both a negative multiplier and an aggregate action below the no-intervention level, y^{no}_{agg}, provided that $\mathbf{d}^{-}_{w}(\mathbf{Z}) \neq \frac{k}{N}\mathbf{1}$. Now, $\Pr[M(\mathbf{Z},N,n,0)<0]>0$ if and only if there exists a configuration $\mathbf{b}(N,n) \in \mathcal{B}(N,n)$ for which $\widehat{f}_{avg}(\mathbf{Z},\mathbf{b},N,n) < \frac{kn}{N}$; the same condition is required for $\Pr\left[Y_{agg}(\mathbf{Z},N,n,0) < y^{ns}_{agg}\right] > 0$. We therefore prove that, for every integer $n \in \{1,\ldots,N-1\}$, min supp $\widehat{F}_{avg}(\mathbf{Z},N,n) < \frac{kn}{N}$.

Order the elements of $\mathbf{d}^{-}_{w}(\mathbf{Z})$ to form the vector $\widetilde{\mathbf{w}}$, with $\widetilde{w}_i \leq \widetilde{w}_{i'}$ if and only if $i \leq i'$. $\mathbf{1}^T\mathbf{d}^{-}_{w}(\mathbf{Z}) = k$, so $\mathbf{1}^T\widetilde{\mathbf{w}} = k$ as well. We want to show that min supp $\widehat{F}_{avg}(\mathbf{Z},N,n) = \sum_{i=1}^{n}\widetilde{w}_i < \frac{kn}{N}$ for every integer $n \in \{1,\ldots,N-1\}$. The proof follows by induction. We first show that $\widetilde{w}_1 < \frac{k}{N}$ and $\widetilde{w}_1+\widetilde{w}_2 < \frac{2k}{N}$; then, given that $\widetilde{w}_1+\cdots+\widetilde{w}_{n-1} < \frac{k(n-1)}{N}$, we prove that $\widetilde{w}_1+\cdots+\widetilde{w}_n < \frac{kn}{N}$ for a general $n \in \{3,\ldots,N-1\}$.

Showing that $\widetilde{w}_1 < \frac{k}{N}$: Suppose that $\widetilde{w}_1 = \frac{k}{N}$. With $\widetilde{w}_i \geq \widetilde{w}_1\ \forall i \in \{2,\ldots,N\}$, $\sum_{i=1}^{N}\widetilde{w}_i = k$ if and only if $\widetilde{w}_1 = \cdots = \widetilde{w}_N = \frac{k}{N}$, which we have ruled out by assumption. Suppose that $\widetilde{w}_1 > \frac{k}{N}$. With $\widetilde{w}_i \geq \widetilde{w}_1\ \forall i \in \{2,\ldots,N\}$, $\sum_{i=1}^{N}\widetilde{w}_i > k$. Since we must have $\sum_{i=1}^{N}\widetilde{w}_i = k$, it follows that $\widetilde{w}_1 < \frac{k}{N}$.

Showing that $\widetilde{w}_1+\widetilde{w}_2 < \frac{2k}{N}$: Since $\widetilde{w}_1 < \frac{k}{N}$, set $\widetilde{w}_1 = \frac{k}{N}-\kappa$ for positive κ. We want $\widetilde{w}_2 < \frac{k}{N}+\kappa$ so that $\widetilde{w}_1+\widetilde{w}_2 < \frac{2k}{N}$. Suppose that $\widetilde{w}_2 \geq \frac{k}{N}+\kappa$. Then $\widetilde{w}_i \geq \frac{k}{N}+\kappa\ \forall i \in \{3,\ldots,N\}$, and $\sum_{i=1}^{N}\widetilde{w}_i = \widetilde{w}_1+\widetilde{w}_2+\sum_{i=3}^{N}\widetilde{w}_i \geq \frac{2k}{N}+(N-2)\left(\frac{k}{N}+\kappa\right) = k+(N-2)\kappa > k$. Since we must have $\sum_{i=1}^{N}\widetilde{w}_i = k$, it follows that $\widetilde{w}_1+\widetilde{w}_2 < \frac{2k}{N}$.

Showing that $\widetilde{w}_1+\cdots+\widetilde{w}_n < \frac{kn}{N}$ for a general integer $n \in \{3,\ldots,N-1\}$:

Suppose that $\widetilde{w}_1 + \cdots + \widetilde{w}_{n-1} < \frac{k(n-1)}{N}$. Set $\widetilde{w}_1 + \cdots + \widetilde{w}_{n-1} = \frac{k(n-1)}{N} - \kappa'$ for positive κ'. We want $\widetilde{w}_n < \frac{k}{N} + \kappa'$. Suppose that $\widetilde{w}_n \geq \frac{k}{N} + \kappa'$. Then $\widetilde{w}_i \geq \frac{k}{N} + \kappa'$ for $i \in \{n+1, \ldots, N\}$ and $\sum_{i=1}^{N} \widetilde{w}_i = \sum_{i=1}^{n-1} \widetilde{w}_i + \sum_{i=n}^{N} \widetilde{w}_i \geq \frac{k(n-1)}{N} - \kappa' + (N-n+1)\left(\frac{k}{N} + \kappa'\right) = k + (N-n)\kappa' > k$. Therefore, $\widetilde{w}_n < \frac{k}{N} + \kappa'$ and $\widetilde{w}_1 + \cdots + \widetilde{w}_n < \frac{kn}{N}$.

Therefore, for every $n \in \{1, \ldots, N-1\}$, provided that $\mathbf{d}_w^-(\mathbf{Z}) \neq \frac{k}{N}\mathbf{1}$, $\min \operatorname{supp} \widehat{F}_{avg}(\mathbf{Z}, N, n) < \frac{kn}{N}$ and we thus have $\Pr[M(\mathbf{Z}, N, n, 0) < 0] = \Pr\left[Y_{agg}(\mathbf{Z}, N, n, 0) < y_{agg}^{ns}\right] > 0$. □

Proof of Proposition 2.9

The proof of Proposition 2.9 follows from the proof of Theorem 1 in Ballester, Calvó-Armengol, and Zenou (2006) coupled with Remarks 2 and 3. □

Proof of Proposition 2.10

In a setting with transfers,

$$\begin{aligned} Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) &= y_{agg}^{no} + \\ &\psi N \epsilon \left[\widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) - \frac{n}{N-n}\left(k - \widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right)\right)\right] \\ &= y_{agg}^{no} + \frac{\psi N^2 \epsilon}{N-n}\left[\widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) - \frac{kn}{N}\right]. \end{aligned}$$

In a setting with stimulus,

$$Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) = y_{agg}^{no} + \psi N \epsilon \widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right).$$

$M\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right) = \frac{dY_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 0\right)}{d\epsilon}$ and $M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) = \frac{dY_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)}{d\epsilon}$. □

Proof of Proposition 2.11

We first define an M-matrix:

Definition B.1 *Matrix* $\mathbf{Z}$ *is an* M-matrix *if it has the form* $\mathbf{Z} = s\mathbf{I} - \mathbf{Q}$*, with* $s > 0$*,* $\mathbf{Q} \geq \mathbf{0}$*, and* $s \geq r(\mathbf{Q})$*. Matrix* $\mathbf{Z}$ *is a* non-singular M-matrix *if* $s > r(\mathbf{Q})$*.*

In environments without strategic substitutes, matrix $-\mathbf{\Sigma}$ is an M-matrix: $s = -\sigma$ and $\mathbf{Q} = \mathbf{\Sigma} - \sigma\mathbf{I}$. Provided that $-\sigma > r(\mathbf{\Sigma} - \sigma\mathbf{I})$, matrix $-\mathbf{\Sigma}$ is a non-singular M-matrix. Then $(-\mathbf{\Sigma})^{-1} \geq \mathbf{0}$. With $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) = \frac{1}{N}\left(-\mathbf{\Sigma}^{-1}\right)^T\mathbf{1}$, it follows that $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) \geq \mathbf{0}$. Given the expressions for $Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ and $M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right)$ in Proposition 2.10, $Y_{agg}\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) \geq y_{agg}^{no}$ with probability 1, and $M\left(-\mathbf{\Sigma}^{-1}, N, n, 1\right) \geq 0$ with probability 1. □

Proof of Proposition 2.12

In the absence of any network-based interaction, $\mathbf{\Sigma} = \sigma\mathbf{I}$, which makes $-\mathbf{\Sigma}^{-1} = -\frac{1}{\sigma}\mathbf{I}$, $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) = -\frac{1}{N\sigma}\mathbf{1}$, and $k = -\frac{1}{\sigma}$. As a result, $\widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) = -\frac{n}{N\sigma} > 0$ with probability 1. Given the expressions for aggregate output and the corresponding economic multiplier in Proposition 2.10, we obtain our result. □

Proof of Proposition 2.13

When $\mathbf{1}^T\mathbf{\Sigma} = \delta\mathbf{1}^T$, $\mathbf{1}^T\mathbf{\Sigma}\mathbf{\Sigma}^{-1} = \delta\mathbf{1}^T\mathbf{\Sigma}^{-1}$, so $\mathbf{1}^T\mathbf{\Sigma}^{-1} = \frac{1}{\delta}\mathbf{1}^T$, and the sum of each column of $\mathbf{\Sigma}^{-1}$ is $\frac{1}{\delta}$. We define $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)$ equal to $-\frac{1}{N}\left(\mathbf{\Sigma}^{-1}\right)^T\mathbf{1}$. Each element of $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right)$ has the same value, and since we set $\mathbf{1}^T\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) = k$, $\mathbf{d}_w^-\left(-\mathbf{\Sigma}^{-1}\right) = \frac{k}{N}\mathbf{1}$ with $k = -\frac{1}{\delta}$. Then, $\widehat{F}_{avg}\left(-\mathbf{\Sigma}^{-1}, N, n\right) = \frac{kn}{N}$ with probability 1, and given the expressions for aggregate output and the corresponding economic multiplier in Proposition 2.10, our result follows. □

Proof of Proposition 2.14

For all $n \in \{1, \ldots, N-1\}$, $Y_{agg}\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, N, n, 0\right) \succeq Y_{agg}\left(-\boldsymbol{\Sigma}^{-1}, N, n, 0\right)$ if $y_{agg}\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n, 0\right) \geq y_{agg}\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n, 0\right)$ for every configuration $\mathbf{b} \in \mathcal{B}(N, n)$. Similarly, for all $n \in \{1, \ldots, N-1\}$, $Y_{agg}\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, N, n, 1\right) \succeq Y_{agg}\left(-\boldsymbol{\Sigma}^{-1}, N, n, 1\right)$ if $y_{agg}\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n, 1\right) \geq y_{agg}\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n, 1\right)$ for every configuration $\mathbf{b} \in \mathcal{B}(N, n)$. Each configuration $\mathbf{b}(N, n)$ of stimulus represents an adjustment to the wealth vector: $\boldsymbol{\omega} \to \boldsymbol{\omega} + \boldsymbol{\rho}$. We therefore wish to show that, for any given vector $\boldsymbol{\omega}' = \boldsymbol{\omega} + \boldsymbol{\rho}$, the aggregate action in the $\boldsymbol{\Sigma}'$ environment exceeds the aggregate action in the $\boldsymbol{\Sigma}$ environment. Define $\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right)$ as the equilibrium vector of agent actions in the $\boldsymbol{\Sigma}'$ environment with wealth vector $\boldsymbol{\omega}'$ and corresponding aggregate action $y_{agg}\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) = \mathbf{1}^T \mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right)$. Similarly define $\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right)$.

Set $\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} + \mathbf{D}$. $[\mathbf{D}]_{ij} \geq 0$ for all pairs (i, j) with at least one strict inequality $[\mathbf{D}]_{ij} > 0$. With $\beta > \lambda r(\mathbf{G})$ and $\beta' > \lambda' r(\mathbf{G}')$, $\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right)$, $\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right) > \mathbf{0}$. We also have that $-\boldsymbol{\Sigma}\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right) = \psi\boldsymbol{\omega}'$ and $-\boldsymbol{\Sigma}'\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) = \psi\boldsymbol{\omega}'$. Decompose $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma} = -\beta\mathbf{I} - \gamma\mathbf{U} + \lambda\mathbf{G}$, and decompose $\boldsymbol{\Sigma}'$ as $\boldsymbol{\Sigma}' = \beta'\mathbf{I} - \gamma'\mathbf{U} + \lambda'\mathbf{G}'$, where $\mathbf{U} = \mathbf{1}\mathbf{1}^T$.

For any vector $\boldsymbol{\omega}'$, $-\left(\boldsymbol{\Sigma}' - \mathbf{D}\right)\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right) = \psi\boldsymbol{\omega}$, and therefore

$$\begin{aligned}\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right)\mathbf{y}^*&\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right) \\ &= \psi\boldsymbol{\omega}' - \gamma'\mathbf{U}\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right) - \mathbf{D}\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}, \mathbf{b}, N, n\right). \quad \text{(B.1)}\end{aligned}$$

We also have $-\boldsymbol{\Sigma}'\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) = \psi\boldsymbol{\omega}'$, and therefore

$$\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right)\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) = \psi\boldsymbol{\omega}' - \gamma'\mathbf{U}\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right). \quad \text{(B.2)}$$

Setting $\gamma' = 0$ and subtracting Equation B.1 from Equation B.2 gives

$$\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right)\left[\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) - \mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n\right)\right]$$
$$= \mathbf{D}\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n\right).$$

$\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right) = \lambda'\left(\frac{\beta'}{\lambda'}\mathbf{I} - \mathbf{G}'\right)$, with $\lambda' > 0$. $\left(\frac{\beta'}{\lambda'}\mathbf{I} - \mathbf{G}'\right)$ is an M-matrix that is nonsingular when $\frac{\beta'}{\lambda'} > \rho\left(\mathbf{G}'\right)$, which immediately follows from the initial assumptions. Therefore, $\left(\frac{\beta'}{\lambda'}\mathbf{I} - \mathbf{G}'\right)$ is inverse-positive, and

$$\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) - \mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n\right)$$
$$= \left(\lambda'\right)^{-1}\left(\frac{\beta'}{\lambda'}\mathbf{I} - \mathbf{G}'\right)^{-1}\mathbf{D}\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n\right) > \mathbf{0},$$

so for any vector of wealth $\boldsymbol{\omega}'$, $\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}, \mathbf{b}, N, n\right) > \mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}, \mathbf{b}, N, n\right)$. □

Proof of Proposition 2.15

Set $\mathbf{\Sigma}' = \mathbf{\Sigma} + \mathbf{D}$, with $[\mathbf{D}]_{ij} \geq 0$ for all pairs (i, j) and at least one strict inequality $[\mathbf{D}]_{ij} > 0$. Also set $\boldsymbol{\omega} = \omega\mathbf{1}$. We have $-\mathbf{\Sigma}\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}\right) = \psi\omega\mathbf{1}$ and $-\mathbf{\Sigma}'\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}\right) = \psi\omega\mathbf{1}$ with $\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}\right), \mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}\right) > \mathbf{0}$ since $\beta > \lambda r\left(\mathbf{G}\right)$ and $\beta' > \lambda' r\left(\mathbf{G}'\right)$. We decompose $\mathbf{\Sigma}$ and $\mathbf{\Sigma}'$ as follows: $\mathbf{\Sigma} = -\beta\mathbf{I} - \gamma\mathbf{U} + \lambda\mathbf{G}$ and $\mathbf{\Sigma}' = -\beta'\mathbf{I} - \gamma'\mathbf{U} + \lambda'\mathbf{G}'$, with $\mathbf{U} = \mathbf{1}\mathbf{1}^T$.

Then, $-\left(\mathbf{\Sigma}' - \mathbf{D}\right)\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}\right) = \psi\omega\mathbf{1}$, and therefore

$$\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right)\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}\right) = \psi\boldsymbol{\omega} - \gamma'\mathbf{U}\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}\right) - \mathbf{D}\mathbf{y}^*\left(-\mathbf{\Sigma}^{-1}\right). \tag{B.3}$$

Meanwhile, $-\mathbf{\Sigma}'\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}\right) = \psi\omega\mathbf{1}$, and therefore,

$$\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right)\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}\right) = \psi\omega\mathbf{1} - \gamma'\mathbf{U}\mathbf{y}^*\left(-\left(\mathbf{\Sigma}'\right)^{-1}\right). \tag{B.4}$$

Setting $\gamma' = 0$ and subtracting Equation B.3 from Equation B.4 gives

$$\left(\beta'\mathbf{I} - \lambda'\mathbf{G}'\right)\left(\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) - \mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right)\right) = \mathbf{D}\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right).$$

It follows that

$$\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) - \mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right) = \left(\lambda'\right)^{-1}\left(\frac{\beta'}{\lambda'}\mathbf{I} - \mathbf{G}'\right)^{-1}\mathbf{D}\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right) > \mathbf{0}$$

since $\lambda' > 0$ and $\left(\frac{\beta'}{\lambda'}\mathbf{I} - \mathbf{G}'\right)$ is an M-matrix that is inverse-positive because $\frac{\beta'}{\lambda'} > r\left(\mathbf{G}'\right)$ by assumption. Thus, $\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) > \mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right)$.

Now, $-\boldsymbol{\Sigma}\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right) = \psi\omega\mathbf{1}$ and $-\boldsymbol{\Sigma}'\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) = \psi\omega\mathbf{1}$. Since $\beta > \lambda r\left(\mathbf{G}\right)$ and $\beta' > \lambda' r\left(\mathbf{G}\right)$,

$$\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right) = \psi\omega\left(-\boldsymbol{\Sigma}^{-1}\right)\mathbf{1} \text{ and } \mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) = \psi\omega\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right)\mathbf{1}.$$

$\boldsymbol{\Sigma}, \boldsymbol{\Sigma}'$ are symmetric, so $\mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right) = \psi\omega\left(-\boldsymbol{\Sigma}^{-1}\right)^T\mathbf{1} = \psi N\omega\mathbf{d}_w^-\left(-\boldsymbol{\Sigma}^{-1}\right)$ and $\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) = \psi\omega\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right)^T\mathbf{1} = \psi N\omega\mathbf{d}_w^-\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right)$. Since $\mathbf{y}^*\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) > \mathbf{y}^*\left(-\boldsymbol{\Sigma}^{-1}\right)$, it immediately follows that $\mathbf{d}_w^-\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}\right) > \mathbf{d}_w^-\left(-\boldsymbol{\Sigma}^{-1}\right)$. Then $\widehat{F}_{avg}\left(-\left(\boldsymbol{\Sigma}'\right)^{-1}, N, n\right) \succeq \widehat{F}_{avg}\left(-\boldsymbol{\Sigma}^{-1}, N, n\right)$ for all $n \in \{1, \ldots, N\}$ and we obtain the result. $\square$

Proof of Proposition 2.16

Agent i's optimization problem is:

$$\max_{y_{i,q}} u_{i,q} = \max_{y_{i,q}} -\sum_{j=1}^{N}\left[\bar{\mathbf{A}}\right]_{ij}\left(y_{i,q} - \left[\mathbf{T}\right]_{ij}\left(y_{j,q-1}\right)\right)^2.$$

From the first-order condition, $\sum_{j=1}^{N}\left[\bar{\mathbf{A}}\right]_{ij} y_{i,q}^* = \sum_{j=1}^{N}\left[\bar{\mathbf{A}}\right]_{ij}\left[\mathbf{T}\right]_{ij}\left(y_{j,q-1}\right)$. By the row-stochasticity of $\bar{\mathbf{A}}$, $y_{i,q}^* = \sum_{j=1}^{N}\left[\bar{\mathbf{A}}\right]_{ij}\left[\mathbf{T}\right]_{ij}\left(y_{j,q-1}\right)$, and so $y_{i,q}^* = \left[\bar{\mathbf{A}} \circ \mathbf{T}\right]_{i*}\mathbf{y}_{q-1}$. It then follows that $\mathbf{y}_q^* = \left(\bar{\mathbf{A}} \circ \mathbf{T}\right)^q\mathbf{y}_0$. $\square$

Proof of Proposition 2.17

Given that $\mathbf{y}_q^* = y\mathbf{1}$ for $q < 0$, the proof follows by induction. For $q = 0$, $\mathbf{y}_q^* = y\mathbf{1} + (\bar{\mathbf{A}} \circ \mathbf{O})^0 \boldsymbol{\rho} = y\mathbf{1} + \boldsymbol{\rho}$. For $q = 1$, given Proposition 2.16, $\mathbf{y}_1^* = (\bar{\mathbf{A}} \circ \mathbf{O})\, \mathbf{y}_0$, so

$$
\begin{aligned}
y_{i,1}^* &= \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij}\, y_{j,0} + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij}\, \mathcal{D}\left(y_{j,0}\right) \\
&= \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij} \left(y + [\boldsymbol{\rho}]_j\right) + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij} \left(2y - \left(y + [\boldsymbol{\rho}]_j\right)\right) \\
&= y + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij}\, [\boldsymbol{\rho}]_j - \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij}\, [\boldsymbol{\rho}]_j \\
&= y + [\bar{\mathbf{A}} \circ \mathbf{O}]_{i*}\, \boldsymbol{\rho},
\end{aligned}
$$

where $[\mathbf{O}]_{ij} = 1$ if $[\mathbf{T}]_{ij} = \mathcal{F}$ and $[\mathbf{O}]_{ij} = -1$ if $[\mathbf{T}]_{ij} = \mathcal{D}$. It then follows that $\mathbf{y}_1^* = y\mathbf{1} + (\bar{\mathbf{A}} \circ \mathbf{O})\, \boldsymbol{\rho}$.

We now assume that $\mathbf{y}_{q-1}^* = y\mathbf{1} + (\bar{\mathbf{A}} \circ \mathbf{O})^{q-1} \boldsymbol{\rho}$, and we demonstrate that

$\mathbf{y}_q^* = y\mathbf{1} + (\bar{\mathbf{A}} \circ \mathbf{O})^q \boldsymbol{\rho}$:

$$\begin{aligned}
y_{i,q}^* &= [\bar{\mathbf{A}} \circ \mathbf{T}]_{i*} \mathbf{y}_{q-1} \\
&= \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij} y_{j,q-1} + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij} \mathcal{D}\left(y_{j,q-1}\right) \\
&= \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij} \left(y + \left[(\bar{\mathbf{A}} \circ \mathbf{O})^{q-1}\right]_{j*} \boldsymbol{\rho}\right) \\
&\qquad + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij} \left(2y - \left(y + \left[(\bar{\mathbf{A}} \circ \mathbf{O})^{q-1}\right]_{j*} \boldsymbol{\rho}\right)\right) \\
&= y + \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{F}}} [\bar{\mathbf{A}}]_{ij} \left[(\bar{\mathbf{A}} \circ \mathbf{O})^{q-1}\right]_{j*} \boldsymbol{\rho} - \sum_{\substack{j \in \{1,\ldots,N\} \\ \text{s.t. } [\mathbf{T}]_{ij} = \mathcal{D}}} [\bar{\mathbf{A}}]_{ij} \left[(\bar{\mathbf{A}} \circ \mathbf{O})^{q-1}\right]_{j*} \boldsymbol{\rho} \\
&= y + [\bar{\mathbf{A}} \circ \mathbf{O}]_{j*} \left(\left[(\bar{\mathbf{A}} \circ \mathbf{O})^{q-1}\right]_{j*} \boldsymbol{\rho}\right) \\
&= y + \left[(\bar{\mathbf{A}} \circ \mathbf{O})^{q}\right]_{j*} \boldsymbol{\rho},
\end{aligned}$$

so $\mathbf{y}_q^* = y\mathbf{1} + (\bar{\mathbf{A}} \circ \mathbf{O})^q \boldsymbol{\rho}$. $\square$

Proof of Proposition 2.18

The period-q aggregate action is:

$$y_{agg,q}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, \mathbf{b}, N, n\right) = y_{agg}^{no} + N\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T \boldsymbol{\rho}.$$

In a setting with positive transfers to n agents and negative transfers to $N-n$ agents, $[\boldsymbol{\rho}]_i = \epsilon$ if agent i is receiving a positive transfer, and $[\boldsymbol{\rho}]_i = -\frac{n\epsilon}{N-n}$ if agent i

is receiving a negative transfer. Therefore,

$$\begin{aligned} Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right) &= y_{agg}^{no} + N\epsilon\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right) \\ &\quad - N\frac{n\epsilon}{N-n}\left(k_q - \widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right)\right) \\ &= y_{agg}^{no} + \frac{N^2\epsilon}{N-n}\left(\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right) - \frac{k_q n}{N}\right) \end{aligned}$$

With $M_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right) = \frac{dY_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right)}{d\epsilon}$ and $IRF_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right) = Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,0\right) - y_{agg}^{no}$, the first part of the Proposition follows.

In a setting with stimulus, $[\boldsymbol{\rho}]_i = \epsilon$ if agent i is receiving stimulus, and $[\boldsymbol{\rho}]_i = 0$ if agent i is not receiving stimulus. Therefore,

$$Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right) = y_{agg}^{no} + N\epsilon\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right).$$

With $M_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right) = \frac{dY_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right)}{d\epsilon}$ and $IRF_q\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right) = Y_{agg,q}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n,1\right) - y_{agg}^{no}$, the second part of the Proposition follows. □

Proof of Proposition 2.19

Set $\bar{\mathbf{A}}\circ\mathbf{O} = \mathbf{I}$. As a result, $\mathbf{d}_w^-\left((\bar{\mathbf{A}}\circ\mathbf{O})^q\right) = \frac{1}{N}\mathbf{1}$ and $k_q = 1$ for all $q \in \mathbb{Z}_+$. Then, $\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right) = \frac{n}{N}$ with probability 1 for all $q \in \mathbb{Z}_+$. The result follows from Proposition 2.18. □

Proof of Proposition 2.20

When $\mathbf{1}^T\left(\bar{\mathbf{A}}\circ\mathbf{O}\right)^q = k_q\mathbf{1}^T$, $\mathbf{d}_w^-\left((\bar{\mathbf{A}}\circ\mathbf{O})^q\right) = \frac{k_q}{N}\mathbf{1}$ and $\widehat{F}_{avg}\left((\bar{\mathbf{A}}\circ\mathbf{O})^q,N,n\right) = \frac{k_q n}{N}$ with probability 1. Proposition 2.20 then follows from Proposition 2.18. □

Proof of Proposition 2.21

If $(\bar{\mathbf{A}}' \circ \mathbf{O}')^q \geq (\bar{\mathbf{A}} \circ \mathbf{O})^q \mathbf{P}$ for some permutation matrix $\mathbf{P}$, then $\left[\mathbf{d}_w^- \left((\bar{\mathbf{A}}' \circ \mathbf{O}')^q\right)\right]^T \geq \left[\mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T \mathbf{P}$ and $\widehat{F}_{avg}\left((\bar{\mathbf{A}}' \circ \mathbf{O}')^q, N, n\right) \succeq \widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)$. Proposition 2.21 then follows from Proposition 2.18. $\square$

Proof of Lemma 2.7

By definition, $\mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right) = \frac{1}{N}\left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]^T \mathbf{1}$ and $k_q = \mathbf{1}^T \mathbf{d}_w^- \left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)$, so $Nk_q = \mathbf{1}^T \left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]^T \mathbf{1} = \sum_{i=1}^N \sum_{j=1}^N \left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]_{ij}$. $k_q \in [-1, 1]$ for all $q \geq 1$ if and only if $\sum_{i=1}^N \sum_{j=1}^N \left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]_{ij} \in [-N, N]$ for all $q \geq 1$. When $q = 1$, $-\bar{\mathbf{A}} \leq \bar{\mathbf{A}} \circ \mathbf{O} \leq \bar{\mathbf{A}}$ element-wise, so $-N = \mathbf{1}^T(-\bar{\mathbf{A}})\mathbf{1} \leq \mathbf{1}^T(\bar{\mathbf{A}} \circ \mathbf{O})\mathbf{1} \leq \mathbf{1}^T\bar{\mathbf{A}}\mathbf{1} = N$, and $\sum_{i=1}^N \sum_{j=1}^N [\bar{\mathbf{A}} \circ \mathbf{O}]_{ij} \in [-N, N]$.

For $q > 1$, $[\bar{\mathbf{A}}^q]_{ij}$, $\left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]_{ij}$, and $-[\bar{\mathbf{A}}^q]_{ij}$ are each the sum of a collection of terms, with each term a product of elements $[\bar{\mathbf{A}}]_{\ell\ell'}$. The collection of terms, ignoring sign, is the same for $[\bar{\mathbf{A}}^q]_{ij}$, $\left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]_{ij}$, and $-[\bar{\mathbf{A}}^q]_{ij}$. For $[\bar{\mathbf{A}}^q]_{ij}$, all of the terms have a positive sign; for $-[\bar{\mathbf{A}}^q]_{ij}$, all of the terms have a negative sign; and for $\left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]_{ij}$, when $\bar{\mathbf{A}} \circ \mathbf{O} \neq \bar{\mathbf{A}}$ and $\bar{\mathbf{A}} \circ \mathbf{O} \neq -\bar{\mathbf{A}}$, the terms have a mixture of positive and negative signs. Therefore, $-\bar{\mathbf{A}}^q \leq (\bar{\mathbf{A}} \circ \mathbf{O})^q \leq \bar{\mathbf{A}}^q$ element-wise. $\bar{\mathbf{A}}$ is row-stochastic, with row stochasticity preserved under matrix multiplication, so $\mathbf{1}^T\bar{\mathbf{A}}^q\mathbf{1} = N$. For $q > 1$, $-N = \mathbf{1}^T(-\bar{\mathbf{A}}^q)\mathbf{1} \leq \mathbf{1}^T(\bar{\mathbf{A}} \circ \mathbf{O})^q\mathbf{1} \leq \mathbf{1}^T\bar{\mathbf{A}}^q\mathbf{1} = N$, and $\sum_{i=1}^N \sum_{j=1}^N \left[(\bar{\mathbf{A}} \circ \mathbf{O})^q\right]_{ij} \in [-N, N]$. $\square$

Proof of Proposition 2.22

From Proposition 2.18, $M_q\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n, 1\right) = N\widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)$. With $k_q \in [-1, 1]$, for every $n \in \{1, \ldots, N-1\}$,

$$\max_{(\bar{\mathbf{A}} \circ \mathbf{O})^q} \left[\max \operatorname{supp} \widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)\right] = 1,$$

which occurs when $\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 \end{pmatrix} \mathbf{P}$ for some permutation matrix $\mathbf{P}$. Meanwhile,

$$\min_{(\bar{\mathbf{A}} \circ \mathbf{O})^q} \left[\min \operatorname{supp} \widehat{F}_{avg}\left((\bar{\mathbf{A}} \circ \mathbf{O})^q, N, n\right)\right] = -1,$$

which occurs when $\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T = \begin{pmatrix} 0 & 0 & \ldots & 0 & -1 \end{pmatrix} \mathbf{P}$ for some permutation matrix $\mathbf{P}$. Consistent with $\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 \end{pmatrix} \mathbf{P}$ is a positive star graph, and consistent with $\left[\mathbf{d}_w^-\left((\bar{\mathbf{A}} \circ \mathbf{O})^q\right)\right]^T = \begin{pmatrix} 0 & 0 & \ldots & 0 & -1 \end{pmatrix} \mathbf{P}$ is a negative star graph. □

Proof of Proposition 2.23

The proof of Proposition 2.23 follows immediately from Theorem 1.1. Provided that $\bar{\mathbf{A}}$ is primitive, $\lim_{q \to \infty} [\bar{\mathbf{A}}^q]_{ij} = \left[\mathbf{w}_\infty^T\right]_j$ for all $i \in \{1, \ldots, N\}$, and therefore $\lim_{q \to \infty} \mathbf{d}_w^-\left(\bar{\mathbf{A}}^q\right) = \mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)$. □

Proof of Proposition 2.24

If $\bar{\mathbf{A}}$ is primitive, $[\bar{\mathbf{A}}]_{ij} > 0$ if and only if $[\bar{\mathbf{A}}]_{ji} > 0$, and all non-zero elements within every row of $\bar{\mathbf{A}}$ have the same value, then by Theorem 1.2, $\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right) = \frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}}$. Since

$\lim_{q\to\infty} \bar{\mathbf{A}}^q = \mathbf{1}\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T$, $\lim_{q\to\infty} \mathbf{d}_w^-\left(\bar{\mathbf{A}}^q\right) = \frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}}$. Setting $\bar{\mathbf{A}} \circ \mathbf{O} = \bar{\mathbf{A}}$,

$$\lim_{q\to\infty} Y_{agg,q}\left(\bar{\mathbf{A}}^q, N, n, 0\right) = y_{agg}^{no} + \frac{N^2\epsilon}{N-n} \lim_{q\to\infty}\left(\widehat{F}_{avg}\left(\bar{\mathbf{A}}^q, N, n\right) - \frac{k_q n}{N}\right).$$

$\lim_{q\to\infty} k_q = \lim_{q\to\infty}\left[\mathbf{d}_{\overline{w}}^-\left(\bar{\mathbf{A}}^q\right)\right]^T \mathbf{1} = 1$. When $n = 1$, $\lim_{q\to\infty} \widehat{F}_{avg}\left(\bar{\mathbf{A}}^q, N, n\right) = \frac{D(\bar{\mathbf{A}})}{\mathbf{1}^T\mathbf{d}}$, so

$$\begin{aligned}\Pr\left[\lim_{q\to\infty} Y_{agg,q}\left(\bar{\mathbf{A}}^q, N, n, 0\right) < y_{agg}^{no}\right] &= \Pr\left[\lim_{q\to\infty} \widehat{F}_{avg}\left(\bar{\mathbf{A}}^q, N, n\right) < \frac{k_q n}{N}\right] \\ &= \Pr\left[\frac{D\left(\bar{\mathbf{A}}\right)}{\mathbf{1}^T\mathbf{d}} < \frac{1}{N}\right] = \Pr\left[D\left(\bar{\mathbf{A}}\right) < \frac{\mathbf{1}^T\mathbf{d}}{N}\right],\end{aligned}$$

and the Proposition follows. □

Proof of Proposition 2.25

We wish to compute $\lim_{q\to\infty}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q$. Graph $\mathcal{G}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right) = \left(\mathcal{V}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right), \mathcal{E}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)\right)$, $\left|\mathcal{V}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)\right| = N$, is structurally balanced, so we can partition $\mathcal{V}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)$ into two disjoint subsets: $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$. Moreover, since $\mathcal{G}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)$ is structurally balanced and $\left|\bar{\mathbf{A}} \circ \mathbf{O}\right| = \bar{\mathbf{A}}$, there exists an $N \times N$ diagonal matrix $\mathbf{\Omega}$ for which $\bar{\mathbf{A}} = \mathbf{\Omega}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)\mathbf{\Omega}$. To make this equality hold, choose an $\ell \in \{1,2\}$ and set $[\mathbf{\Omega}]_{ii} = 1$ for $\forall i \in \mathcal{V}_\ell$ and $[\mathbf{\Omega}]_{ii} = -1$ for $\forall i \in \mathcal{V}_{-\ell}$. Matrix $\mathbf{\Omega}$ equals its inverse, $\mathbf{\Omega}^{-1}$, so $\bar{\mathbf{A}}^q = \left(\mathbf{\Omega}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)\mathbf{\Omega}\right)^q = \mathbf{\Omega}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q\mathbf{\Omega}$ and $\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q = \mathbf{\Omega}\bar{\mathbf{A}}^q\mathbf{\Omega}$. Therefore,

$$\lim_{q\to\infty}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q = \lim_{q\to\infty} \mathbf{\Omega}\bar{\mathbf{A}}^q\mathbf{\Omega} = \mathbf{\Omega}\left(\lim_{q\to\infty} \bar{\mathbf{A}}^q\right)\mathbf{\Omega} = \mathbf{\Omega}\mathbf{1}\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T\mathbf{\Omega},$$

where $\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T\bar{\mathbf{A}} = \left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T$. Specifically, $\left[\lim_{q\to\infty}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q\right]_{ij} = \left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]_j$ if $i, j \in \mathcal{V}_\ell$ and $\left[\lim_{q\to\infty}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q\right]_{ij} = -\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]_j$ if $i \in \mathcal{V}_\ell$ and $j \in \mathcal{V}_{-\ell}$ for $\ell \in \{1,2\}$. Equivalently, $\lim_{q\to\infty}\left(\bar{\mathbf{A}} \circ \mathbf{O}\right)^q = \left(\mathbf{1}\left[\mathbf{w}_\infty\left(\bar{\mathbf{A}}\right)\right]^T\right) \circ \mathbf{O}$, where $[\mathbf{O}]_{ij} = 1$ if $i, j \in \mathcal{V}_\ell$ and $[\mathbf{O}]_{ij} = -1$ if $i \in \mathcal{V}_\ell$ and $j \in \mathcal{V}_{-\ell}$ for $\ell \in \{1,2\}$. □

Proof of Proposition 2.26

Solving the representative consumer's problem:

$$\max_{c_1,\ldots,c_N} \prod_{i=1}^{N} c_i^{\eta_i} \quad \text{s.t.} \quad \sum_{i=1}^{N} p_i c_i = w \sum_{i=1}^{N} \ell_i + \sum_{i=1}^{N} \pi_i.$$

Setting $\sum_{i=1}^{N} \ell_i = 1$, the first-order condition is:

$$p_i c_i = \eta_i \left(w + \sum_{i=1}^{N} \pi_i \right). \tag{B.5}$$

Solving the firm's optimization problem:

$$\max_{x_{1i},\ldots,x_{Ni},\ell_i} p_i x_i - \sum_{j=1}^{N} p_j x_{ji} - w\ell_i \quad \text{s.t.} \quad x_i = A_i^{\alpha_i} \ell_i^{\alpha_i} \left(\prod_{j=1}^{N} x_{ji}^{[\mathbf{\Lambda}]_{ji}} \right)^{\beta_i}.$$

Rewriting, we have:

$$\max_{x_{1i},\ldots,x_{Ni},\ell_i} p_i A_i^{\alpha_i} \ell_i^{\alpha_i} \left(\prod_{j=1}^{N} x_{ji}^{[\mathbf{\Lambda}]_{ji}} \right)^{\beta_i} - \sum_{j=1}^{N} p_j x_{ji} - w\ell_i.$$

The first-order conditions are:

$$w = \frac{\alpha_i p_i x_i}{\ell_i} \tag{B.6}$$

and

$$p_j = \frac{[\mathbf{\Lambda}]_{ji}\, \beta_i p_i x_i}{x_{ji}}. \tag{B.7}$$

The profit of each firm is:

$$\pi_i = p_i x_i - \sum_{j=1}^{N} p_j x_{ji} - w\ell_i.$$

From Equations B.6 and B.7,

$$\pi_i = p_i x_i - \sum_{j=1}^{N} [\mathbf{\Lambda}]_{ji}\, \beta_i p_i x_i - \alpha_i p_i x_i.$$

Matrix $\mathbf{\Lambda}$ is column-stochastic, so

$$\pi_i = (1 - \alpha_i - \beta_i)\, p_i x_i = 0. \tag{B.8}$$

From Equations B.5 and B.8,

$$c_i = \frac{\eta_i w}{p_i}, \tag{B.9}$$

and from Equation B.7,

$$x_{ij} = \frac{[\mathbf{\Lambda}]_{ij}\, \beta_j p_j x_j}{p_i}. \tag{B.10}$$

Substituting the expressions for c_i and x_{ij} in Equations B.9 and B.10 into the goods market clearing condition, $x_i = c_i + \sum_{j=1}^{N} x_{ij}$, we have:

$$p_i x_i = \eta_i w + \sum_{j=1}^{N} [\mathbf{\Lambda}]_{ij}\, \beta_j p_j x_j.$$

Define $y_j = p_j x_j$. Then $\mathbf{y} = \boldsymbol{\eta} w + \mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta})\, \mathbf{y}$.

Lemma B.2 $\mathbf{I} - \mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta})$ *is invertible.*

Proof. $\mathbf{I} - \mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta})$ is an M-matrix. It is therefore invertible if $1 > r\,(\mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta}))$. We know that any matrix-induced norm satisfies the inequality $\|\mathbf{Z}\| > |\mu|$ for any matrix $\mathbf{Z}$ with eigenvalue μ. We take the infinity norm of matrix $(\mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta}))^T$:

$$\left\| (\mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta}))^T \right\|_\infty = \max_{i \in \{1,\ldots,N\}} \sum_{j=1}^{N} \left[(\mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta}))^T \right]_{ij} = \max_{i \in \{1,\ldots,N\}} \beta_i$$

when $\mathbf{\Lambda}$ is column-stochastic. Since $\max_{i \in \{1,\ldots,N\}} \beta_i \in (0, 1)$, $r\,(\mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta})) < 1$, and $\mathbf{I} - \mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta})$ is invertible. ∎

Following Lemma B.2, we obtain $\mathbf{y}^* = (\mathbf{I} - \mathbf{\Lambda}\, \mathrm{diag}\,(\boldsymbol{\beta}))^{-1}\, \boldsymbol{\eta} w$. □

Proof of Proposition 2.27

The proof of Proposition 2.27 immediately follows from the proof of Proposition 2.26.

□

Proof of Proposition 2.28

In a setting with transfers,

$$Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,0\right)=y_{agg}^{no}+N\epsilon\left[\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n\right)\right.$$
$$\left.-\frac{n}{N-n}\left(k-\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n\right)\right)\right],$$

so

$$Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,0\right)=y_{agg}^{no}$$
$$+\frac{N^2\epsilon}{N-n}\left[\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n\right)-\frac{kn}{N}\right]$$

with $M\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,0\right)=\frac{dY_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,0\right)}{d\epsilon}$.

In a setting with stimulus,

$$Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,1\right)=y_{agg}^{no}+N\epsilon\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n\right)$$

and $M\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,1\right)=\frac{dY_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\operatorname{diag}(\boldsymbol{\beta}))^{-1},N,n,1\right)}{d\epsilon}$. □

Proof of Proposition 2.29

From the expressions for $Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right)$ and $M\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right)$ we see that

$$\Pr\left[Y_{agg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right)\geq y_{agg}^{no}\right] = \Pr\left[M\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1},N,n,1\right)\geq 0\right]=1$$

for every $n\in\{1,\ldots,N-1\}$ if and only if

$$\Pr\left[\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1},N,n\right)\geq 0\right]=1.$$

Now, $\Pr\left[\widehat{F}_{avg}\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1},N,n\right)\geq 0\right]=1$ for every $n\in\{1,\ldots,N-1\}$ if and only if $\mathbf{d}_w^-\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}\right)\geq\mathbf{0}$. Since $\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta})$ is an M-matrix with $1>r\,(\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))$, $\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta})$ is inverse-positive which makes $\mathbf{d}_w^-\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}\right)\geq\mathbf{0}$. □

Proof of Proposition 2.30

In general, the vector of average weighted in-degrees is

$$\left[\mathbf{d}_w^-\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}\right)\right]^T=\frac{1}{N}\mathbf{1}^T\,(\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}.$$

Setting $\mathbf{\Lambda}=\mathbf{I}$, $\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta})=\mathrm{diag}\,(\mathbf{1}-\boldsymbol{\beta})$ and $(\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}=\mathrm{diag}\begin{pmatrix}\frac{1}{1-\beta_1} & \cdots & \frac{1}{1-\beta_N}\end{pmatrix}$; the expression for $\mathbf{d}_w^-\left((\mathbf{I}-\mathbf{\Lambda}\,\mathrm{diag}\,(\boldsymbol{\beta}))^{-1}\right)$ then follows. □

Proof of Proposition 2.31

We first demonstrate the following: If $\beta_1=\cdots=\beta_N\equiv\beta$, then both GDP and the corresponding economic multiplier are invariant to configuration. Suppose that $\beta_1 = \cdots = \beta_N \equiv \beta$. Then $\text{diag}(\boldsymbol{\beta}) = \beta\mathbf{I}$. $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta})) = \mathbf{1}^T(\mathbf{I} - \beta\boldsymbol{\Lambda}) = (1-\beta)\mathbf{1}^T$, so $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1} = (1-\beta)\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}$ and $\frac{1}{1-\beta}\mathbf{1}^T = \mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}$. Therefore, $\left[\mathbf{d}_w^-\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}\right)\right]^T = \frac{1}{N}\frac{1}{1-\beta}\mathbf{1}^T$ and $k = \frac{1}{1-\beta}$. $\widehat{F}_{avg}\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n\right) = \frac{n}{N}\frac{1}{1-\beta}$ with probability 1, and we obtain the expressions for $Y_{agg}\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 0\right)$, $M\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 0\right)$, $Y_{agg}\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 1\right)$, and $M\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 1\right)$.

We next demonstrate the following: If both GDP and the corresponding economic multiplier are invariant to configuration, then $\beta_1 = \cdots = \beta_N \equiv \beta$. Suppose that both GDP and the corresponding economic multiplier are invariant to configuration. If $Y_{agg}\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 0\right)$, $M\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 0\right)$, $Y_{agg}\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 1\right)$, and $M\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n, 1\right)$ are invariant to configuration, then $\widehat{F}_{avg}\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}, N, n\right)$ is invariant to configuration. We must then have $\mathbf{d}_w^-\left((\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}\right) = \frac{k}{N}\mathbf{1}$, or equivalently, $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1} = k\mathbf{1}^T$. It follows that $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))^{-1}(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta})) = k\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta}))$, so $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta})) = \frac{1}{k}\mathbf{1}^T$. In general, $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta})) = \begin{pmatrix} 1-\beta_1 & \cdots & 1-\beta_N \end{pmatrix}$ by the column-stochasticity of $\boldsymbol{\Lambda}$. For $\mathbf{1}^T(\mathbf{I} - \boldsymbol{\Lambda}\,\text{diag}(\boldsymbol{\beta})) = \frac{1}{k}\mathbf{1}^T$, we must have $\beta_1 = \cdots = \beta_N \equiv \beta$, and we then have $N(1-\beta) = Nk$, or equivalently, $k = \frac{1}{1-\beta}$. $\square$

Appendix C

Appendix to Chapter 3

C.1 Proofs

Proof of Proposition 3.1

$E\Pi_i(\mathbf{w}_i, L, \ell) = E\left[\mathbf{w}_i^T \delta \mathbf{B}(L, \ell)\right]$, where $\mathbf{B}(L, \ell)$ is a random vector whose elements are $B_j \sim \text{Bern}\left(\frac{\ell}{L}\right)$, $j \in \{1, \ldots, L\}$ and $\mathbf{1}^T\mathbf{w}_i = k_i$. Therefore,

$$E\Pi_i(\mathbf{w}_i, L, \ell) = \delta E\left([\mathbf{w}_i]_1 B_1 + \cdots + [\mathbf{w}_i]_L B_L\right) = \delta \sum_{j=1}^{L} [\mathbf{w}_i]_j EB_j = \delta k_i \frac{\ell}{L}.$$

$E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) = E\left[\mathbf{w}_{agg}^T \delta \mathbf{B}(L, \ell)\right]$, where $\mathbf{1}^T\mathbf{w}_{agg} = k_{agg}$. Therefore,

$$\begin{aligned} E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) = \delta E\left([\mathbf{w}_{agg}]_1 B_1 + \cdots + [\mathbf{w}_{agg}]_L B_L\right) &= \delta \sum_{j=1}^{L} [\mathbf{w}_{agg}]_j EB_j \\ &= \delta k_{agg} \frac{\ell}{L}. \quad \square \end{aligned}$$

Proof of Proposition 3.2

$\mathrm{Var}\,\Pi_i\left(\mathbf{w}_i, L, \ell\right) = \mathrm{Var}\left[\mathbf{w}_i^T\,\delta\,\mathbf{B}\left(L,\ell\right)\right]$, where $\mathbf{B}\left(L,\ell\right)$ is a random vector whose elements are $B_j \sim \mathrm{Bern}\left(\frac{\ell}{L}\right)$, $j \in \{1,\ldots,L\}$. The random variables $B_j \sim \mathrm{Bern}\left(\frac{\ell}{L}\right)$ are identically distributed but not independent. We then have

$$\mathrm{Var}\,\Pi_i\left(\mathbf{w}_i, L, \ell\right) = \delta^2\,\mathrm{Var}\left[\mathbf{w}_i^T\mathbf{B}\left(L,\ell\right)\right] = \delta^2\frac{\ell}{L}\left(1-\frac{\ell}{L}\right)\frac{L}{L-1}L\,\mathrm{Var}\,W_i$$

from Theorem 1.9 in Chapter 1 and Lemma 2.2 in Chapter 2. Similarly, with

$$\mathrm{Var}\,\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \mathrm{Var}\left[\mathbf{w}_{agg}^T\,\delta\,\mathbf{B}\left(L,\ell\right)\right] = \delta^2\,\mathrm{Var}\left[\mathbf{w}_{agg}^T\mathbf{B}\left(L,\ell\right)\right],$$

we then have $\mathrm{Var}\,\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \delta^2\frac{\ell}{L}\left(1-\frac{\ell}{L}\right)\frac{L}{L-1}L\,\mathrm{Var}\,W_{agg}$. $\square$

Proof of Proposition 3.3

$$\pi_i\left(\mathbf{w}_i, \boldsymbol{\epsilon}, L, \ell\right) = \mathbf{w}_i^T\boldsymbol{\epsilon}\left(L,\ell\right) = \delta \sum_{\substack{j\in\{1,\ldots,L\}\\ \text{s.t. } [\boldsymbol{\epsilon}]_j=\delta}} \left[\mathbf{w}_i\right]_j \quad \text{and}$$

$$\pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell\right) = \mathbf{w}_{agg}^T\boldsymbol{\epsilon}\left(L,\ell\right) = \delta \sum_{\substack{j\in\{1,\ldots,L\}\\ \text{s.t. } [\boldsymbol{\epsilon}]_j=\delta}} \left[\mathbf{w}_{agg}\right]_j.$$

With $\delta < 0$, the Proposition then follows. $\square$

Proof of Proposition 3.4

$\Pi_i\left(\mathbf{w}_i, L, \ell\right) = \delta X\left(\mathbf{w}_i, L, \ell\right)$, where random variable $X\left(\mathbf{w}_i, L, \ell\right)$ has realizations $x\left(\mathbf{w}_i, \mathbf{b}, L, \ell\right) = \mathbf{w}_i^T\mathbf{b}\left(L,\ell\right)$, the elements of $\mathbf{b}\left(L,\ell\right)$ are either 0 or 1, and $\mathbf{1}^T\mathbf{b}\left(L,\ell\right) = \ell$. From Theorem 1.13 in Chapter 1, provided that condition (c) holds, and after re-labelling some variables, we have that

$$\left|G_{\frac{X\left(\mathbf{w}_i,L,\ell\right)-EX\left(\mathbf{w}_i,L,\ell\right)}{\left(\mathrm{Var}\,X\left(\mathbf{w}_i,L,\ell\right)\right)^{1/2}}}(t) - J\left(\widehat{\mathbf{w}}, L, \ell, t\right)\right| < C_4 \times \sum_{j=1}^{L}\left|\widehat{w}_j\right|^5$$

with $\widehat{w}_j = \frac{[\mathbf{w}_i]_j - EW_i}{\sqrt{L \operatorname{Var} W_i}}$. Now, $G_{\frac{\Pi_i(\mathbf{w}_i,L,\ell)-E\Pi_i(\mathbf{w}_i,L,\ell)}{\left(\operatorname{Var}\Pi_i(\mathbf{w}_i,L,\ell)\right)^{1/2}}}(t) = G_{\frac{X(\mathbf{w}_i,L,\ell)-EX(\mathbf{w}_i,L,\ell)}{\left(\operatorname{Var}X(\mathbf{w}_i,L,\ell)\right)^{1/2}}}(t)$, so the Proposition holds for the individual financial institutions $i \in \{1, \ldots, M\}$. Similarly, $\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \delta X\left(\mathbf{w}_{agg}, L, \ell\right)$, where random variable $X\left(\mathbf{w}_{agg}, L, \ell\right)$ has realizations $x\left(\mathbf{w}_{agg}, \mathbf{b}, L, \ell\right) = \mathbf{w}_{agg}^T \mathbf{b}\left(L, \ell\right)$, the elements of $\mathbf{b}\left(L, \ell\right)$ are either 0 or 1, and $\mathbf{1}^T \mathbf{b}\left(L, \ell\right) = \ell$. Given Theorem 1.13 in Chapter 1, Proposition 3.4 therefore also holds when approximating the CDF for the entire financial system. □

Proof of Lemma 3.1

$$
\begin{aligned}
\mathbf{w}_i^T\left(\boldsymbol{\epsilon}\left(L,\ell\right)\circ\bar{\mathbf{p}}\right) &= \begin{pmatrix}[\mathbf{w}_i]_1 & [\mathbf{w}_i]_2 & \cdots & [\mathbf{w}_i]_L\end{pmatrix}\widehat{\delta}\begin{pmatrix}[\bar{\mathbf{p}}]_1\\ [\bar{\mathbf{p}}]_2\\ \vdots\\ [\bar{\mathbf{p}}]_L\end{pmatrix}\circ\begin{pmatrix}[\mathbf{b}\left(L,\ell\right)]_1\\ [\mathbf{b}\left(L,\ell\right)]_2\\ \vdots\\ [\mathbf{b}\left(L,\ell\right)]_L\end{pmatrix}\\
&= \begin{pmatrix}[\mathbf{w}_i]_1 & \frac{[\mathbf{w}_i]_2[\bar{\mathbf{p}}]_2}{[\bar{\mathbf{p}}]_1} & \cdots & \frac{[\mathbf{w}_i]_L[\bar{\mathbf{p}}]_L}{[\bar{\mathbf{p}}]_1}\end{pmatrix}\begin{pmatrix}\widehat{\delta}\,[\bar{\mathbf{p}}]_1\\ \widehat{\delta}\,[\bar{\mathbf{p}}]_1\\ \vdots\\ \widehat{\delta}\,[\bar{\mathbf{p}}]_1\end{pmatrix}\circ\begin{pmatrix}[\mathbf{b}\left(L,\ell\right)]_1\\ [\mathbf{b}\left(L,\ell\right)]_2\\ \vdots\\ [\mathbf{b}\left(L,\ell\right)]_L\end{pmatrix}\\
&= \left(\widehat{\delta}\,[\bar{\mathbf{p}}]_1\right)\mathbf{v}_i^T\left(\mathbf{1}_{L\times 1}\circ\mathbf{b}\left(L,\ell\right)\right)\\
&= \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,\mathbf{v}_i^T\mathbf{b}\left(L,\ell\right).
\end{aligned}
$$

To show that $\mathbf{w}_{agg}^T\left(\boldsymbol{\epsilon}\left(L,\ell\right)\circ\bar{\mathbf{p}}\right) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,\mathbf{v}_{agg}^T\mathbf{b}\left(L,\ell\right)$, simply replace $\mathbf{w}_i$ with $\mathbf{w}_{agg}$ in the above derivation. □

Proof of Proposition 3.5

$E\Pi_i(\mathbf{w}_i, L, \ell) = E\left[\widehat{\delta}\,[\bar{\mathbf{p}}]_1\,\mathbf{v}_i^T\mathbf{B}(L,\ell)\right]$, where $\mathbf{B}(L,\ell)$ is a random vector whose elements are $B_j \sim \text{Bern}\left(\frac{\ell}{L}\right)$, $j \in \{1,\ldots,L\}$, and $\mathbf{1}^T\mathbf{v}_i = k_i$. Therefore,

$$E\Pi_i(\mathbf{w}_i, L, \ell) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,E\left([\mathbf{v}_i]_1 B_1 + \cdots + [\mathbf{v}_i]_L B_L\right) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1 \sum_{j=1}^{L} [\mathbf{v}_i]_j\,EB_j = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,k_i\frac{\ell}{L}.$$

Meanwhile, $E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) = E\left[\widehat{\delta}\,[\bar{\mathbf{p}}]_1\,\mathbf{v}_{agg}^T\mathbf{B}(L,\ell)\right]$, where $\mathbf{1}^T\mathbf{v}_{agg} = k_{agg}$. Then,

$$\begin{aligned} E\Pi_{agg}(\mathbf{w}_{agg}, L, \ell) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,E\left([\mathbf{v}_{agg}]_1 B_1 + \cdots + [\mathbf{v}_{agg}]_L B_L\right) &= \widehat{\delta}\,[\bar{\mathbf{p}}]_1 \sum_{j=1}^{L} [\mathbf{v}_{agg}]_j\,EB_j \\ &= \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,k_{agg}\frac{\ell}{L}. \quad \square \end{aligned}$$

Proof of Proposition 3.6

This proposition immediately follows from Proposition 3.2. Simply substitute $\mathbf{w}_i$ for $\mathbf{v}_i$, $\mathbf{w}_{agg}$ for $\mathbf{v}_{agg}$, and δ for $\widehat{\delta}\,[\bar{\mathbf{p}}]_1$. $\square$

Proof of Proposition 3.7

$$\pi_i(\mathbf{w}_i, \boldsymbol{\epsilon}, L, \ell) = \mathbf{w}_i^T(\boldsymbol{\epsilon}(L,\ell)\circ\bar{\mathbf{p}}) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,\mathbf{v}_i^T\mathbf{b}(L,\ell) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1 \sum_{\substack{j\in\{1,\ldots,L\} \\ \text{s.t. } [\mathbf{b}]_j=1}} [\mathbf{v}_i]_j, \text{ and}$$

$$\pi_{agg}(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, L, \ell) = \mathbf{w}_{agg}^T(\boldsymbol{\epsilon}(L,\ell)\circ\bar{\mathbf{p}}) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,\mathbf{v}_{agg}^T\mathbf{b}(L,\ell) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1 \sum_{\substack{j\in\{1,\ldots,L\} \\ \text{s.t. } [\mathbf{b}]_j=1}} [\mathbf{v}_{agg}]_j.$$

With $\widehat{\delta} < 0$, the Proposition then follows. $\square$

Proof of Proposition 3.8

This proposition follows from Proposition 3.4. Set $\Pi_i(\mathbf{w}_i, L, \ell) = \widehat{\delta}\,[\bar{\mathbf{p}}]_1\,X(\mathbf{v}_i, L, \ell)$, where random variable $X(\mathbf{v}_i, L, \ell)$ has realizations $x(\mathbf{v}_i, \mathbf{b}, L, \ell) = \mathbf{v}_i^T\mathbf{b}(L,\ell)$. Similarly, set $\Pi_{agg}\left(\mathbf{w}_{agg}, L, \ell\right) = \widehat{\delta}\left[\bar{\mathbf{p}}\right]_1 X\left(\mathbf{v}_{agg}, L, \ell\right)$, where random variable $X\left(\mathbf{v}_{agg}, L, \ell\right)$ has realizations $x\left(\mathbf{v}_{agg}, \mathbf{b}, L, \ell\right) = \mathbf{v}_{agg}^T \mathbf{b}\left(L, \ell\right)$. □

Proof of Proposition 3.9

First method of proof: Define $\left\{\sigma_j\left(\delta\right)\right\}_{j=1}^{L!}$ as the family of all possible permutations of the elements in δ. Then,

$$\begin{aligned}
E\Pi_i\left(\mathbf{w}_i, \delta\right) &= \frac{1}{L!}\sum_{j=1}^{L!} \mathbf{w}_i^T \sigma_j\left(\delta\right) \\
&= \frac{1}{L!}\sum_{j=1}^{L!}\sum_{m=1}^{L} \left[\mathbf{w}_i\right]_m \left[\sigma_j\left(\delta\right)\right]_m \\
&= \sum_{m=1}^{L} \left[\mathbf{w}_i\right]_m \left[\frac{1}{L!}\sum_{j=1}^{L!} \left[\sigma_j\left(\delta\right)\right]_m\right] \\
&= \sum_{m=1}^{L} \left[\mathbf{w}_i\right]_m \left[\frac{1}{L!}\sum_{n=1}^{L} \left(L-1\right)! \left[\delta\right]_n\right] \\
&= \sum_{m=1}^{L} \left[\mathbf{w}_i\right]_m \left[\frac{1}{L}\sum_{n=1}^{L} \left[\delta\right]_n\right] \\
&= \left(\frac{\mathbf{1}^T\delta}{L}\right) k_i.
\end{aligned}$$

Substituting $\mathbf{w}_i$ for $\mathbf{w}_{agg}$, $E\Pi_{agg}\left(\mathbf{w}_{agg}, \delta\right) = \left(\frac{\mathbf{1}^T\delta}{L}\right) k_{agg}$. □

Second method of proof: $E\Pi_i\left(\mathbf{w}_i, \delta\right) = E\left[\mathbf{w}_i^T \mathbf{\Delta}\right]$, where $\mathbf{\Delta}$ is an $L \times 1$ random vector whose elements are Δ_j, $j \in \{1, \ldots, L\}$. The random variables $\Delta_1, \ldots, \Delta_L$ are identically distributed, with the corresponding CDF $G_{\Delta_j}(t) = \frac{1}{L}\sum_{m=1}^{L} \mathbb{1}_{[\delta]_m \leq t}$. Then,

$$\begin{aligned} E\Pi_i(\mathbf{w}_i, \delta) &= E\left([\mathbf{w}_i]_1 \Delta_1 + \cdots + [\mathbf{w}_i]_L \Delta_L\right) \\ &= \sum_{j=1}^{L} [\mathbf{w}_i]_j E\Delta_j \\ &= \left(\frac{\mathbf{1}^T\delta}{L}\right) \sum_{j=1}^{L} [\mathbf{w}_i]_j \\ &= \left(\frac{\mathbf{1}^T\delta}{L}\right) k_i. \end{aligned}$$

Substituting $\mathbf{w}_i$ for $\mathbf{w}_{agg}$, $E\Pi_{agg}(\mathbf{w}_{agg}, \delta) = \left(\frac{\mathbf{1}^T\delta}{L}\right) k_{agg}$. $\square$

Proof of Proposition 3.10

$$\begin{aligned} &\operatorname{Var}\Pi_{agg}(\mathbf{w}_{agg}, \delta) = \operatorname{Var}\left(\mathbf{w}_{agg}^T \boldsymbol{\Delta}\right) \\ &= \operatorname{Var}\left(\left[\mathbf{w}_{agg}\right]_1 \Delta_1 + \cdots + \left[\mathbf{w}_{agg}\right]_L \Delta_L\right) \\ &= \sum_{j=1}^{L} \operatorname{Var}\left(\left[\mathbf{w}_{agg}\right]_j \Delta_j\right) + \sum_{j=1}^{L} \sum_{\substack{m=1 \\ j \neq m}}^{L} \operatorname{Cov}\left(\left[\mathbf{w}_{agg}\right]_j \Delta_j, \left[\mathbf{w}_{agg}\right]_m \Delta_m\right) \\ &= \sum_{j=1}^{L} \left(\left[\mathbf{w}_{agg}\right]_j\right)^2 \operatorname{Var}\Delta_j + \sum_{j=1}^{L} \sum_{\substack{m=1 \\ j \neq m}}^{L} \left[\mathbf{w}_{agg}\right]_j \left[\mathbf{w}_{agg}\right]_m \operatorname{Cov}\left(\Delta_j, \Delta_m\right) \\ &= \left(\operatorname{Var}\Delta_j\right) \sum_{j=1}^{L} \left(\left[\mathbf{w}_{agg}\right]_j\right)^2 + \left(E\left[\Delta_j \Delta_m\right] - \left(E\Delta_j\right)\left(E\Delta_m\right)\right) \sum_{j=1}^{L} \sum_{\substack{m=1 \\ j \neq m}}^{L} \left[\mathbf{w}_{agg}\right]_j \left[\mathbf{w}_{agg}\right]_m. \end{aligned}$$

We would like to find a simple expression to replace

$\sum_{j=1}^{L}\sum_{\substack{m=1\\ j\neq m}}^{L}\left[\mathbf{w}_{agg}\right]_j\left[\mathbf{w}_{agg}\right]_m$. Note the following:

$$\begin{aligned}
\operatorname{Var} W_{agg} &= \frac{1}{L}\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j - \frac{\sum_{m=1}^{L}\left[\mathbf{w}_{agg}\right]_m}{L}\right)^2 \\
&= \frac{1}{L}\left(\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 - \frac{1}{L}\left(\sum_{j=1}^{L}\left[\mathbf{w}_{agg}\right]_j\right)^2\right) \\
&= \frac{1}{L}\left(\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 - \frac{1}{L}\left(\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 + \sum_{j=1}^{L}\sum_{\substack{m=1\\ j\neq m}}^{L}\left[\mathbf{w}_{agg}\right]_j\left[\mathbf{w}_{agg}\right]_m\right)\right) \\
&= \frac{1}{L}\frac{L-1}{L}\left[\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 - \frac{1}{L-1}\sum_{j=1}^{L}\sum_{\substack{m=1\\ j\neq m}}^{L}\left[\mathbf{w}_{agg}\right]_j\left[\mathbf{w}_{agg}\right]_m\right].
\end{aligned}$$

Therefore,

$$\sum_{j=1}^{L}\sum_{\substack{m=1\\ j\neq m}}^{L}\left[\mathbf{w}_{agg}\right]_j\left[\mathbf{w}_{agg}\right]_m = (L-1)\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 - L^2\operatorname{Var} W_{agg},$$

and

$$\begin{aligned}
\operatorname{Var}\Pi_{agg}\left(\mathbf{w}_{agg},\delta\right) &= \left(\operatorname{Var}\Delta_j\right)\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 \\
&\quad + \left(E\left[\Delta_j\Delta_m\right] - \left(E\Delta_j\right)\left(E\Delta_m\right)\right)\left[(L-1)\sum_{j=1}^{L}\left(\left[\mathbf{w}_{agg}\right]_j\right)^2 - L^2\operatorname{Var} W_{agg}\right].
\end{aligned}$$

To compute $\operatorname{Var}\Pi_i\left(\mathbf{w}_i,\delta\right)$ for $i\in\{1,\ldots,M\}$, replace $\mathbf{w}_{agg}$ with $\mathbf{w}_i$, and replace W_{agg} with W_i. □

Proof of Proposition 3.11

$\pi_i\left(\mathbf{w}_i,\boldsymbol{\epsilon},\delta\right) = \mathbf{w}_i^T\boldsymbol{\epsilon}\left(\delta\right)$ and $\pi_{agg}\left(\mathbf{w}_{agg},\boldsymbol{\epsilon},\delta\right) = \mathbf{w}_{agg}^T\boldsymbol{\epsilon}\left(\delta\right)$. The elements in the multisets $\left\{\tilde{w}_j\right\}_{j=1}^{L}$, $\left\{\tilde{x}_j\right\}_{j=1}^{L}$, and $\left\{\tilde{\delta}_j\right\}_{j=1}^{L}$ are real-valued and weakly increasing with index. By the rearrangement inequality,

$$\tilde{w}_L\tilde{\delta}_1 + \tilde{w}_{L-1}\tilde{\delta}_2 + \cdots + \tilde{w}_1\tilde{\delta}_L \\ \leq \tilde{w}_{\sigma(1)}\tilde{\delta}_1 + \tilde{w}_{\sigma(2)}\tilde{\delta}_2 + \cdots + \tilde{w}_{\sigma(L)}\tilde{\delta}_L \leq \tilde{w}_1\tilde{\delta}_1 + \tilde{w}_2\tilde{\delta}_2 + \cdots + \tilde{w}_L\tilde{\delta}_L$$

and

$$\tilde{x}_L\tilde{\delta}_1 + \tilde{x}_{L-1}\tilde{\delta}_2 + \cdots + \tilde{x}_1\tilde{\delta}_L \\ \leq \tilde{x}_{\sigma(1)}\tilde{\delta}_1 + \tilde{x}_{\sigma(2)}\tilde{\delta}_2 + \cdots + \tilde{x}_{\sigma(L)}\tilde{\delta}_L \leq \tilde{x}_1\tilde{\delta}_1 + \tilde{x}_2\tilde{\delta}_2 + \cdots + \tilde{x}_L\tilde{\delta}_L$$

for all possible permutations $\tilde{w}_{\sigma(1)}, \tilde{w}_{\sigma(2)}, \ldots, \tilde{w}_{\sigma(L)}$ of elements $\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_L$, and for all possible permutations $\tilde{x}_{\sigma(1)}, \tilde{x}_{\sigma(2)}, \ldots, \tilde{x}_{\sigma(L)}$ of elements $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L$. $\square$

Proof of Lemma 3.2

$$\mathbf{w}_i^T \left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right) = \begin{pmatrix} [\mathbf{w}_i]_1 & [\mathbf{w}_i]_2 & \cdots & [\mathbf{w}_i]_L \end{pmatrix} \begin{pmatrix} \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_1 \\ \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_2 \\ \vdots \\ \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_L \end{pmatrix} \circ \begin{pmatrix} [\bar{\mathbf{p}}]_1 \\ [\bar{\mathbf{p}}]_2 \\ \vdots \\ [\bar{\mathbf{p}}]_L \end{pmatrix}$$

$$= \begin{pmatrix} [\mathbf{w}_i]_1 & [\mathbf{w}_i]_2 & \cdots & [\mathbf{w}_i]_L \end{pmatrix} \begin{pmatrix} [\bar{\mathbf{p}}]_1 \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_1 \\ [\bar{\mathbf{p}}]_2 \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_2 \\ \vdots \\ [\bar{\mathbf{p}}]_L \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_L \end{pmatrix}$$

$$= \begin{pmatrix} [\mathbf{w}_i]_1 & \frac{[\mathbf{w}_i]_2 [\bar{\mathbf{p}}]_2}{[\bar{\mathbf{p}}]_1} & \cdots & \frac{[\mathbf{w}_i]_L [\bar{\mathbf{p}}]_L}{[\bar{\mathbf{p}}]_1} \end{pmatrix} \begin{pmatrix} \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_1 \\ \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_2 \\ \vdots \\ \left[\boldsymbol{\epsilon}\left(\widehat{\delta}\right)\right]_L \end{pmatrix} [\bar{\mathbf{p}}]_1$$

$$= [\bar{\mathbf{p}}]_1 \, \mathbf{v}_i^T \boldsymbol{\epsilon}\left(\widehat{\delta}\right).$$

To show that $\mathbf{w}_{agg}^T \left(\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \circ \bar{\mathbf{p}}\right) = [\bar{\mathbf{p}}]_1 \, \mathbf{v}_{agg}^T \boldsymbol{\epsilon}\left(\widehat{\delta}\right)$, simply replace $\mathbf{w}_i$ with $\mathbf{w}_{agg}$ in the above derivation. □

Proof of Proposition 3.12

Following the first method of proof in Proposition 3.9, define $\left\{\sigma_j\left(\widehat{\delta}\right)\right\}_{j=1}^{L!}$ as the family of all possible permutations of the elements in $\widehat{\delta}$. Then,

$$\begin{aligned} E\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right) &= \frac{1}{L!}\sum_{j=1}^{L!}[\bar{\mathbf{p}}]_1\,\mathbf{v}_i^T\sigma_j\left(\widehat{\delta}\right) \\ &= [\bar{\mathbf{p}}]_1\,\frac{1}{L!}\sum_{j=1}^{L!}\mathbf{v}_i^T\sigma_j\left(\widehat{\delta}\right) \\ &= [\bar{\mathbf{p}}]_1\left(\frac{\mathbf{1}^T\widehat{\delta}}{L}\right)k_i. \end{aligned}$$

Substituting $\mathbf{w}_i$ for $\mathbf{w}_{agg}$, $E\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\left(\frac{\mathbf{1}^T\widehat{\delta}}{L}\right)k_{agg}$. □

Proof of Proposition 3.13

$$\operatorname{Var}\Pi_i\left(\mathbf{w}_i, \widehat{\delta}\right) = \operatorname{Var}\left([\bar{\mathbf{p}}]_1\,\mathbf{v}_i^T\widehat{\boldsymbol{\Delta}}\right) = \left([\bar{\mathbf{p}}]_1\right)^2\operatorname{Var}\left(\mathbf{v}_i^T\widehat{\boldsymbol{\Delta}}\right)$$

and

$$\operatorname{Var}\Pi_{agg}\left(\mathbf{w}_{agg}, \widehat{\delta}\right) = \operatorname{Var}\left([\bar{\mathbf{p}}]_1\,\mathbf{v}_{agg}^T\widehat{\boldsymbol{\Delta}}\right) = \left([\bar{\mathbf{p}}]_1\right)^2\operatorname{Var}\left(\mathbf{v}_{agg}^T\widehat{\boldsymbol{\Delta}}\right).$$

To compute $\operatorname{Var}\left(\mathbf{v}_i^T\widehat{\boldsymbol{\Delta}}\right)$, we substitute $\mathbf{w}_i$ with $\mathbf{v}_i$ and $\boldsymbol{\Delta}$ with $\widehat{\boldsymbol{\Delta}}$ in Proposition 3.10. To compute $\operatorname{Var}\left(\mathbf{v}_{agg}^T\widehat{\boldsymbol{\Delta}}\right)$, we substitute $\mathbf{w}_{agg}$ with $\mathbf{v}_{agg}$ and $\boldsymbol{\Delta}$ with $\widehat{\boldsymbol{\Delta}}$ in Proposition 3.10. This Proposition then follows. □

Proof of Proposition 3.14

For all $i \in \{1, \ldots, M\}$,

$$\pi_i\left(\mathbf{w}_i, \boldsymbol{\epsilon}, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\,\mathbf{v}_i^T\boldsymbol{\epsilon}\left(\widehat{\delta}\right) \text{ and } \pi_{agg}\left(\mathbf{w}_{agg}, \boldsymbol{\epsilon}, \widehat{\delta}\right) = [\bar{\mathbf{p}}]_1\,\mathbf{v}_{agg}^T\boldsymbol{\epsilon}\left(\widehat{\delta}\right).$$

The elements in the multisets $\{\tilde{v}_j\}_{j=1}^{L}$, $\{\tilde{x}_j\}_{j=1}^{L}$, and $\{\tilde{\delta}_j\}_{j=1}^{L}$ are real-valued and weakly increasing with index. This Proposition then follows from the rearrangement inequality, as detailed in Proposition 3.11. □