Gas, weed, and fumes: Three essays in empirical environmental economics

By

Edward A. Rubin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Maximilian Auffhammer, Chair
Professor Meredith Fowlie
Professor W. Reed Walker

Summer 2018

Abstract

Gas, Weed, and Fumes: Three essays in empirical environmental economics

by

Edward A. Rubin

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Maximilian Auffhammer, Chair

This dissertation presents a three-part study in modern empirical environmental economics. In these three studies, I focus on five core economic issues—equity, incentives, environmental quality, consumer behavior, and causality—and ask what environmental economics can teach us about three common topics: energy consumption, cannabis legalization, and pesticide application.

The first chapter examines how residential natural gas consumers respond to changes in the price of natural gas. With 70 million consumers, residential natural gas has grown to a first-order policy issue. This first chapter provides the first causally identified, microdata-based estimates of residential natural-gas demand elasticities using a panel of 300 million bills in California. To overcome multiple sources of endogeneity, we employ a two-pronged strategy: we interact (1) a spatial discontinuity along the service areas of two major natural-gas utilities with (2) an instrumental-variables strategy using the utilities' differing rules/behaviors for internalizing upstream spot-market prices. We then demonstrate substantial seasonal and income-based heterogeneities underly this elasticity. These heterogeneities suggest unexplored policies that are potentially efficiency-enhancing and pro-poor.

The second chapter explores what may be unintended—or unconsidered—results of cannabis legalization. Cannabis legalization advocates often argue that cannabis legalization offers the potential to reduce the private and social costs related to criminalization and incarceration—particularly for marginalized populations. While this assertion is theoretically plausible, it boils down to an empirically testable hypothesis that remains untested: does legalizing a previously illegal substance (cannabis) reduce arrests, citations, and general law-enforcement contact? The second chapter of this dissertation provides the first causal evidence that cannabis legalization need not necessarily reduce criminalization—and under the right circumstances, may in fact increase police incidents/arrests for both cannabis products and non-cannabis drugs. First, I present a theoretical model of police effort and drug consumption that demonstrates the importance of substitution and incentives for this hypothesis. I then empirically show that before legalization, drug-incident trends in Denver, Colorado matched trends in many other US cities. However, following cannabis legalization in Colorado, drug incidents spike sharply in Denver, while trends in

comparison cities (unaffected by Colorado's legalization) remain stable. This spike in drug-related police incidents occurs both for cannabis and non-cannabis drugs. Synthetic-control and difference-in-differences empirical designs corroborate the size and significance of this empirical observation, estimating that Colorado's legalization of recreational cannabis nearly doubled police-involved drug incidents in Denver. This chapter's results present important lessons for evaluating the effects and equity of policies ranging legalization to criminal prosecution to policing.

Finally, the third chapter investigates the roles pesticides play in local air quality. Many policymakers, public-health advocates, and citizen groups question whether current pesticide regulations properly equate the marginal social costs of pesticide applications to their marginal social benefits—with particular concern for negative health effects stemming from pesticide exposure. Additionally, recent research and policies in public health, epidemiology, and economics emphasize how fine particulate matter (PM2.5) concentrations harm humans through increased mortality, morbidity, mental health issues, and a host of socioeconomic outcomes. This chapter presents the first empirical evidence that aerially applied pesticides increase local PM2.5 concentrations. To causally estimate this effect, I combine the universe of aerial pesticide applications in the five southern counties of California's San Joaquin Valley (1.8M reports) with the U.S. EPA's PM2.5 monitoring network—exploiting spatiotemporal variation in aerial pesticide applications and variation in local wind patterns. I find significant evidence that (upwind) aerial pesticide applications within 1.5km increase local PM2.5 concentrations. The magnitudes of the point estimates suggest that the top decile of aerial applications may sufficiently increase local PM2.5 to warrant concern for human health.

Jointly, the three parts of this dissertation aim to carefully administer causally minded econometrics, in conjunction with environmental economic theory, to answer unresolved, policy-relevant questions.

For Michelle,
There is no sufficient way to say thank you.
You are incredible. I'm just a computer.

For our family,
Your support, examples, and love
got us here and keep us going.

For the kids,
You make our lives so much brighter.

C | Contents

F | Figures

T | Tables

Acknowledgements

Coauthor, data, and funding Chapter 1 of this dissertation is coauthored with Professor Maximilian Auffhammer of UC Berkeley. I thank the Pacific Gas and Electric Company and the Southern California Gas Company for making their billing data available. I also thank and acknowledge the Giannini Foundation for funding research contained Chapter 3. All errors are my own.

Personal thanks I am immensely grateful to Maximilian Auffhammer for his guidance, help, and kindness. I deeply appreciate the attention, support, and mentoring I've received from Reed Walker. I thank Meredith Fowlie and Solomon Hsiang for their time, insights, and encouragement. Thank you all.

1 | Natural gas price elasticities and optimal cost recovery under consumer heterogeneity: Evidence from 300 million natural gas bills

Chapter abstract With 70 million consumers, residential natural gas has grown to a first-order policy issue. This paper provides the first causally identified, microdata-based estimates of residential natural-gas demand elasticities using a panel of 300 million bills in California. To overcome multiple sources of endogeneity, we employ a two-pronged strategy: we interact (1) a spatial discontinuity along the service areas of two major natural-gas utilities with (2) an instrumental-variables strategy using the utilities' differing rules/behaviors for internalizing upstream spot-market prices. We then demonstrate substantial seasonal and income-based heterogeneities underly this elasticity. These heterogeneities suggest unexplored policies that are potentially efficiency-enhancing and pro-poor.

1.1 Introduction

Coal dominated all other fossil fuels throughout the late 19^{th} and most of the 20^{th} centuries and powered unprecedented economic transformation in the United States and many other major economies. The recent arrival of a new technology enabling gas extraction from below the surface, hydraulic fracturing ("frac(k)ing"), is unearthing ample supplies of low-cost natural gas that will foreseeably fuel the first half of the 21st century. Fracking received significant exemptions from the Clean Air Act, the Clean Water Act, and the Safe Drinking Water Act via the Energy Policy Act of 2005 (Environmental Protection Agency 2013), potentially furthering the rise of natural gas within energy markets. Natural gas production in the United States has expanded dramatically, and natural gas prices have fallen considerably, often residing at half of their pre-2005 levels (Hausman and Kellogg 2015). In 2016, natural gas surpassed coal as the main source of energy for electricity generation in the United States and half of US residences used natural gas as their main heating fuel (U.S. Energy Information Administration 2016b). US residential consumers, depending on the severity of the winter, spend 50-80 billion dollars per year on natural gas (BLS, 2017). The average household spends about as much money on natural gas as it spends on water (BLS, 2017).

The low price and newly abundant volumes of natural gas, coupled with natural gas's status as the cleanest and most efficient fossil fuel (Levine, Carpenter, and Thapa 2014; National Academy of Sciences 2016), have prompted broad public and policy support for the use of this fuel both in end uses and in the generation of electricity.[1] Such support partially stems from natural gas's low carbon content per BTU, leading some to refer to natural gas as a "bridge fuel," *bridging* society toward a future powered by largely carbon-free sources of renewable energy.

Natural gas is not without critics. The most common criticisms of current natural gas policy center on environmental degradation, including groundwater contamination, the possible triggering of small earthquakes, increases in air pollution, and higher incidence of accidents from the large number of trucks servicing fracking sites (Glanz 2009; Bao and Eaton 2016). More broadly, researchers have critiqued inefficient and potentially regressive pricing (and regulatory) regimes used in the consumer-facing side of the industry (Borenstein and Davis 2012; Davis and Muehlegger 2010).

Despite its policy relevance, there is a relative dearth of (well) identified estimates for the own-price elasticity of the demand for natural gas.[2] Specifically, we are unable to find any published research that pairs consumer-level data with appropriate identification strategies to causally estimate a price elasticity of demand for natural gas that carries a causal interpretation. Table 1.1 lists the past studies, the type of data used, and the resulting estimates of the own-price elasticity of demand. As Table 1.1 shows, past papers either estimate the elasticity of demand for residential natural gas using aggregated data (*e.g.*, Hausman and Kellogg 2015; Davis and Muehlegger 2010) or using micro data with average prices (*e.g.*, Alberini, Gans, and Velez-Lopez 2011; Meier and Rehdanz 2010).[3] The majority of these papers do not attempt to deal with bias resulting from multiple sources of simultaneity, which we discuss below.

Research on the price elasticity of demand for natural gas faces two major challenges: insufficient data and multiple potential sources of endogeneity. Many of the available datasets aggregate households' consumption across both space and time. This aggregation—coupled with utilities' multi-tiered volumetric pricing regimes,

[1]The fact that an increasingly large share of natural gas is produced in the United States also wins natural gas considerable political support (Levine, Carpenter, and Thapa 2014).

[2]Though several previous papers have offered estimates for the price elasticity of demand for residential natural gas, the existing natural-gas demand elasticity literature addressing these issues is sparse relative to that of the electricity literature (Rehdanz 2007). A cursory Google Scholar search returns approximately 148,000 results related to *economics, elasticities, and electricity*; equivalent searches for *coal* and *gasoline* return approximately 70,000 results each. A similar search for articles related to *natural gas* finds fewer than 40,000 results. (The authors performed these searches in January 2017.)

[3]The exception is Rehdanz 2007, who uses a two-period sample from West Germany, where it appears average price equalled marginal price. Rehdanz does not, however, address the endogeneity of price.

income-based discounts, and fixed charges—makes it impossible for researchers to match consumers to the actual prices they face. Aggregation across customers and seasons also inhibits research into heterogeneity across consumers. Perhaps most importantly, research on the elasticity of demand for natural gas must also consider multiple potential sources of endogeneity. The first source of endogeneity is the classic simultaneity that stems from the fact that quantity and price result from the equilibrium in a system of equations. Unlike the electricity sector, for natural-gas customers' rates change on a monthly basis—updating as a function of gas wholesale prices paid by the retail utilities. The second source of endogeneity results from the fact that price is mechanically a function of quantity in a block-rate price regime. As a household's consumption increases, its marginal price increases in discrete steps; consequently, average price also increases with consumption. Thus, a simple, unidentified regression of quantity on price will result in an incorrect—and potentially positive—estimate of the price elasticity of demand.

This paper is the first to address these aggregation and endogeneity issues so as to causally identify the elasticity of demand for residential natural gas. In order to overcome both the aggregation problem and the endogeneity issues due to increased block rate pricing, the paper uses a dataset of approximately 300 million residential natural gas bills in California and builds on Ito 2014 to exploit a spatial discontinuity based upon the boundary between the service areas of two large natural gas utilities.

The paper makes four concrete contributions to the literature on estimating price elasticities of demand.

First, the natural gas market provides a unique setting with monthly pass-through of wholesale prices to consumers.[4] To overcome potential endogeneity, we combine a spatial discontinuity with a *supply-shifting* instrumental variables (IV) approach. We instrument the utilities' consumer-facing prices with the weekly average spot price of natural gas at a major natural gas distribution hub in Louisiana (the *Henry Hub*). This instrument is valid, as we know the formula of how utilities pass-through the price (providing a strong first stage), and the price is determined prior to within-bill consumption (strengthening the exclusion restriction). Jointly, the spatial discontinuity and spot-price instrument isolate plausibly exogenous variation in residential natural gas prices between neighboring households due to the two utilities' differential pass-through of spot-market prices—and due to households' arbitrarily different billing-period start dates. In other words, as a result of this two-part empirical strategy, coupled with the rich set of fixed effects that our dataset allows, the identifying variation in the residential price of natural gas comes from (**1**) on which side of a long-established border between utility-owned natural gas networks the household is located, (**2**) the subtly different pricing rules and buying strategies governing the two rate-of-return-earning utilities as they individually respond to price variation in the natural gas spot market, and (**3**) on which day of the month a household's bill begins.

[4]Electricity markets generally update cost pass-through with less frequency and with lower variance.

Our second contribution builds upon the fact that we observe whether households are part of a low-income program that provides subsidized natural gas to households. We use this knowledge to estimate price elasticities by high- versus low-income households. Third, we observe billing at a roughly monthly frequency for the households in our sample, allowing us to estimate seasonal elasticities, as well as price elasticities specific to income-group by season. Finally, due both to the temporal resolution of the data and to the fact that we observe households over long periods of time in the same housing structure, we determine whether households respond to current or lagged prices. This result bears evidence on the salience of natural gas bills.

We find that on average, the price elasticity of demand for residential natural gas ranges from −0.23 to −0.17. Importantly, we find evidence of heterogeneity in this elasticity along the dimensions of season and income. Both lower-income and higher-income households are essentially inelastic to price in summer months. However, in winter months, lower-income households are substantially more elastic to price than higher-income households. We discuss unexplored policies with the potential to increase both efficiency and progressivity in settings where externalities from natural gas consumption are priced. Finally, we show evidence that households respond to lagged electricity prices—a result consistent with rational inattention following from the difficulties households face in finding real-time information on natural gas consumption and prices. In addition to motivating previously unexplored policies with the potential to enhance efficiency and reduce the burden on the poor, these heterogeneity findings also supply insights into other *pooled* elasticity estimates that do not consider underlying heterogeneity.

1.2 Institutional setting

In order to identify a causal estimate of the price elasticity of natural gas demand, we need to explain the institutional and physical setup of the natural gas industry in the United States. This market is commonly divided into four segments: (1) production and processing, (2) transportation, (3) storage, and (4) local distribution companies (LDCs). Figure 1.1 illustrates the basic institutional organization of the natural gas industry.[5] The four segments we discuss below roughly follow Figure 1.1 except that they exclude end users (those users who only consume natural gas) and the liquid natural gas import/export-based segments of the market. While this paper focuses on the behavior of residential natural gas consumers, part of our identification strategy relies upon a basic understanding of the wider industry, specifically in understanding which instruments may shift supply without affecting demand. After discussing these four segments, we then describe the multi-tier pric-

[5]We include liquid natural gas (LNG) in the figure for completeness, but liquid natural gas does not play a large role in the natural gas market in the United States: LNG imports currently account for less than one percent of natural gas imports and accounted for three percent of imports at their peak in 2007 (Levine, Carpenter, and Thapa 2014). For this reason, we omit LNG for the rest of this paper.

ing structure employed by the two Californian natural gas utilities discussed in this paper.

1.2.1 Market segments

Production and processing Natural gas enters the market at the wellhead where it is produced and first sold (Brown and Yücel 1993). Some wells produce only natural gas, while other wells produce natural gas in addition to crude oil (Levine, Carpenter, and Thapa 2014). The raw product then moves from wellheads to processors. Processors remove impurities and separate the raw product into multiple commodities (separating "natural gas" from "natural gas liquids") (Levine, Carpenter, and Thapa 2014).

Transportation High-pressure pipelines transport processed natural gas from production and processing areas to both intermediate users (storage facilities, processors, LDCs) and final users (electricity generators, industrial users, commercial users, and residential users). Figure 1.2 maps this pipeline network for the continental United States. Private companies own and operate segments of the pipeline; these pipeline companies' rates are regulated at the state level and the national level (Levine, Carpenter, and Thapa 2014). Extensive spot markets and futures markets exist for natural gas. Louisiana's Henry Hub connects to 13 intrastate and interstate pipelines. The Henry Hub is the designated delivery point for the New York Mercantile Exchange's natural gas futures contracts and the Henry Hub price is generally regarded as a nationally relevant price (Levine, Carpenter, and Thapa 2014). Figure 1.3 depicts the Henry Hub spot price from 1997 through 2016. Transportation costs represent a substantial percentage of natural gas prices; according to Levine, Carpenter, and Thapa, in 2011–2012, 72 percent of consumers' average heating costs originated in "transmission and distribution charges".[6] Thus, the natural gas transportation network creates a nationally integrated market and simultaneously contributes to a sizable portion of the prices paid by natural gas end users.

Storage Storage plays a major role in several parts of the natural gas market, but all parties store mainly for the same reason: volatility within the market. Due to its major roles in heating and electricity production, natural gas demand is strongly driven by weather and can be unpredictable in the short run. To combat price volatility and to be able to meet peak demand, both local distribution companies and large natural gas consumers store gas underground (Levine, Carpenter, and Thapa 2014). Producers utilize storage to smooth production.

Local distribution companies Local distribution companies' primary function is distributing natural gas to their contracted end users—industrial, residential, and

[6]Levine, Carpenter, and Thapa also note that in 2007–2008 "transmission and distribution charges" accounted for 41 percent of consumers' average heating costs. It is worth keeping in mind that consumers' average heating costs fell approximately 20 percent in this period.

commercial consumers of natural gas. To accomplish this task, LDCs purchase natural gas through both spot markets and longer-term contracts. In addition, LDCs own and operate their own pipeline and storage networks. To cover the fixed costs involved in their pipelines, storage, and administration, LDCs often utilize a combination of two-part tariffs and multi-tiered pricing regimes—though some utilities fold all of their costs into their volumetric pricing. State utility commissions (*e.g.*, the California Public Utilities Commission) regulate LDCs' price regimes, allowing the LDCs to earn a regulated rate of return (Brown and Yücel 1993; Davis and Muehlegger 2010; Levine, Carpenter, and Thapa 2014).

1.2.2 Natural gas pricing in California

The California Public Utilities Commission (CPUC) regulates both of the utilities from which we draw data in this paper: Pacific Gas and Electric Company (PG&E) and Southern California Gas Company (SoCalGas). Because this paper analyzes residential natural gas consumers' responses to natural-gas retail prices, the most relevant regulations facing PG&E and SoCalGas are CPUC's price and quantity regulations. In addition, the California Energy Commission (CEC) defines geographic climate zones (see Figure A.1), which, in part, determine households' price schedules (California Energy Commission 2015, 2017).

For PG&E's and SoCalGas's residential consumers, a household's bill depends upon five elements:[7]

1. The **two-tiered price schedule** set by the utility
2. The **total volume of natural gas consumed** during the billing period
3. The **season** (*summer* or *winter*) in which the bill occurs
4. The **climate zone** into which the household's physical location falls
5. The household's **CARE (California Alternate Rates for Energy) status**[8]

Figure A.2 provides an example of a typical residential natural gas bill from PG&E.

Both PG&E and SoCalGas utilize two-tiered pricing regimes. The California Energy Commission divides California into 16 climate zones in which households' needs for heating should be relatively homogeneous (California Energy Commission 2015, 2017; Pacific Gas and Electric Company 2016). The utilities also divide the year into heating (winter) and non-heating (summer) seasons. Based upon a household's climate zone (determined by the household's location) and the season,

[7]Consumers' billing periods do not perfectly align with calendar months. However, PG&E's and SoCalGas's price changes do align with calendar months (during the years that our data cover). The two utilities deal with this misalignment of billing periods and price regimes slightly differently. PG&E calculates individual bills for each calendar month under the assumption that consumption is constant throughout the billing period. SoCalGas calculates a single bill using time-weighted average prices (averaging across the different price regimes). These methods are equivalent under a single linear price but differ under the actual multi-tiered price regimes. Please see the *Calculating bills* section in the appendix for more detail.

[8]The previously mentioned program that provides subsidized energy rates to low-income households in California.

the CPUC determines a volume of natural gas that should be adequate for heating during the course of one day. This volume of natural gas is called the household's *daily allowance*. Multiplying the household's *daily allowance* by the number of days in the billing period gives the household's *total allowance* for the bill. For each unit (*therm*[9]) of natural gas up to the bill's *total allowance*, the household pays the first tier's per-unit price (*baseline price*). For each unit of gas above the household's *total allowance*, the household pays the second tier's per-unit price (*excess price*). Figure 1.4a illustrates an example of the two-tier block-pricing regime used by PG&E and SoCalGas. Figure 1.4b depicts how residential consumers' (daily) tier-one allowances vary through time within a given climate zone (PG&E's climate zone *R* and SoCalGas's climate zone *1*). Figure A.1 depicts California's 16 California Energy Commission (CEC) defined climate zones.

Each month, the utilities update their price schedules. The absolute difference between the first-tier price and the second-tier price also varies but tends to remain constant for several months.[10] These monthly price changes allow the utilities to charge customers at rates that reflect the prevailing price of natural gas. In fact, the utilities tie their price updates to their costs—thus linking residential rates to spot market prices.[11] If the utilities wish to change the way in which their prices are tied to market prices and other costs, they must receive authorization following a review process with CPUC. Figure 1.6a illustrates these monthly price-regime changes and the fairly fixed step between the two tiers. Figure 1.6b depicts the correlation between the utilities' baseline (first-tier) prices and the spot market price of natural gas at the Henry Hub.

A household's participation in the CARE (California Alternate Rates for Energy) program also affects the prices that the household faces. Households qualify for CARE by either meeting low-income qualifications or receiving benefits from one of several state or federal assistance programs (*e.g.*, Medi-Cal or the National School Lunch Program) (Southern California Gas Company 2016). CARE prices are 80 percent of standard prices at both tiers. In addition to giving us the household's correct pricing regime, we use CARE status to identify low-income households.

1.3 Data

1.3.1 Natural gas billing data

The billing data in this paper come from two major utilities in California: Pacific Gas and Electric Company (PG&E) and Southern California Gas Company (SoCalGas). The PG&E data cover residential natural gas bills in PG&E's territory from January 2003 through December 2014. The SoCalGas data cover residential natural gas bills

[9]The utilities in this paper work in units of volume called *therms*. One therm is equal to 100,000 Btu (U.S. Energy Information Administration 2016c).

[10]The utilities differ in the frequencies at which which they change this absolute difference: PG&E adjusts the distance between the two tiers' price much more frequently than SoCalGas.

[11]The utilities report their *weighted average costs of gas* to the CPUC.

from May 2010 through September 2015. Thus, the two utilities' data overlap from May 2010 through December 2014. After excluding zip codes with fewer than 50 households, PG&E's service area covers 597 5-digit zip codes (680,846 9-digit zip codes) with a total of 5,888,276 households and 180,663,705 bills. After excluding zip codes with fewer than 50 households, SoCalGas's service area covers 611 5-digit zip codes (610,207 9-digit zip codes) with a total of 2,526,503 households and 95,335,393 bills. The left side of Figure 1.5 depicts PG&E's and SoCalGas's service areas at the 5-digit zip code level. Table 1.2 provides a brief summary of the billing data with regard to the numbers of bills, households, zip codes, and monetary values of the bills. Tables 1.2 and 1.3 summarize prices, quantities, and other variables of interest—pooling across all observations and also splitting the data by season or CARE status. Both tables summarize the full dataset—all zip codes across both utilities–and a subset of the data based upon all 5-digit zip codes served by both utilities. We describe this subset in detail below in the *Empirical strategy* section.

The utilities' billing data are at the household-bill level: a single row of the dataset represents a single billing period for a given household. Table A.17 describes the variables (columns) in this dataset. We follow the natural gas utilities' convention in defining a household (or customer) as the interaction between a unique utility account and a unique physical location identifier.

We also utilize historical data on pricing from the two utilities. As described above, these pricing data include (1) each utility's monthly two-tier pricing regime and (2) the daily allowance for each climate zone during each season. After joining these pricing data to the households' billing data, we are able to determine both the marginal price and average price (and average marginal price) for each bill received by each household. We forgo analyses below the five-digit zip code because (1) many households are missing their full 9-digit zip codes (the *plus-four codes* are missing), (2) many of the 9-digit zip codes do not match into ZIP4 databases, and (3) our identification strategy already utilizes within-zip-code variation (discussed in detail below).

1.3.2 Weather data

Data on daily weather observations originate from the PRISM project at Oregon State University (PRISM Climate Group 2004). We match this local, daily weather data to the household consumption data at the day by 5-digit-zip-code level. The PRISM dataset contains daily gridded maximum and minimum temperature for the continental United States at a grid cell resolution of roughly 2.5 miles (4 km). Figure A.3 maps a single day of average temperature from the PRISM data for the continental United States. We observe these daily data for California from 1980–2015. In order to match the weather grids to zip codes, we obtained a GIS layer of zip codes from ESRI (Esri 2017), which is based on US Postal Service delivery routes for 2013. For small zip codes not identified by the shape file, we purchased the location of these zip codes from a private vendor[12]. We matched the PRISM grids to the zip code shapes and averaged the daily temperature data across multiple grids within each zip code for each day. For zip codes identified as a point, we use the daily weather observation in the grid at that point. This exercise results in a complete daily record of minimum and maximum temperatures—as well as precipitation—at the zip-code level from 1980–2015.

1.4 Empirical strategy

In this section we describe the empirical strategy that we use to identify the price elasticity of demand for residential natural gas consumers. First, we present the basic estimating equation that motivates the paper's results. Next, we discuss the inherent challenges to identification in this setting. We then discuss potential solutions to these challenges and detail which of these solutions are feasible in this paper's specific setting. Finally, before moving to the results, we provide evidence for the validity of the instruments.

1.4.1 Estimating equation

The relationship at the heart of this paper's elasticity estimates is

$$\log(q_{i,t}) = \eta \log(p_{i,t}) + \lambda_{i,t} + \varepsilon_{i,t} \tag{1.1}$$

where i and t index household and time; q denotes quantity demanded; and p denotes price. Rather than choosing a specific type of price, we present results for five variants of price. These five types of price include the price that classical economic theory deems relevant—the marginal price—in addition to average price, average marginal price, baseline (first-tier) price, and *simulated* marginal price (defined and discussed below).[13] In the results section, we also discuss which lag of price is most salient to consumers (see Figure 1.7 for an example and a brief discussion of price lags). The term $\lambda_{i,t}$ represents household fixed effects, time-based fixed effects, and/or household-by-time fixed effects—depending on the specification. Our main specification in this paper uses household fixed effects and city by month-of-sample fixed effects (*e.g.*, Fresno in January 2010; also called city by year by month). A causally identified estimate of η yields the own-price elasticity of demand.

1.4.2 Challenges to identification

Two main sources of endogeneity threaten identification in equation 1.1.

The first challenge in identifying this own-price elasticity of demand is the potential endogeneity that results from the simultaneous determination of price

[12] zip-codes.com

[13] We define *average marginal price* as the quantity-weighted marginal price paid by a customer during her billing period. *Average marginal* price does not include fixed charges, while *average* price does.

and quantity that results from the equilibrium of supply and demand—*simultaneity* (*e.g.*, Woolridge 2009). In the presence of simultaneity, standard ordinary least squares (OLS) fails to properly treat the endogeneity inherent in (1.1). As discussed above, many papers in the natural gas literature ignore this potential source of bias while estimating the price elasticity of demand—relying upon fixed effects, uncorrelated demand and supply shocks, and/or assumptions of exogenous prices. If simultaneity is indeed present in this setting, then the estimates in these papers will recover biased estimates for the elasticity of demand for residential natural gas.

A second challenge to identification in this paper results from our paper's specific context: the two-tiered price schedule within California's natural gas market. Put simply, in tiered pricing regimes, the marginal price is a (weakly increasing, monotonic) function of quantity. For the same reason, average price is also a function of quantity. Thus, when a household consumes more, its marginal and average prices mechanically increase. In terms of identifying the price elasticity of demand, this is *bad variation*: the marginal price that a household faces is endogenous because the marginal price is correlated with unobserved demand shocks (Ito 2014). This bias is a specific form of simultaneity often called *reverse causality*.

In practice, one generally cannot sign the bias resulting from the classical simultaneity of price and quantity without making further assumptions regarding the correlation of supply and demand shocks. On the other hand, the bias resulting from marginal and average prices being a function of quantity results in upwardly biased estimates of demand *elasticities*. In extreme cases, this latter case of bias can yield estimates that suggest upward-sloping demand curves.

Table 1.5 demonstrates the consequences of failing to address these challenges to identification by estimating the price elasticity of demand—η in equation 1.1 via ordinary least squares (OLS) using marginal price (columns 1–3) and baseline (first-tier) price (columns 4–6). We also vary the set of controls for each price. For a given price, the leftmost columns apply the simplest set of controls. The "identification strategy" presented in Table 1.5 makes no attempt to correct for the aforementioned potential biases outside of a fairly rich set of fixed effects—household fixed effects and city by month-of-sample fixed effects. Each regression controls for within-bill heating degree days (HDDs) during the billing period.[14] The leftmost column for each price uses a five percent sample of all bills from PG&E and SoCalGas (sampled at the five-digit zip code). The remaining columns (columns 2, 3, 4, and 5) use a border-discontinuity motivated sample in which we keep all zip codes where the zip code receives natural gas from both PG&E and SoCalGas (discussed in detail below; also see Figure 1.5). The leftmost and center columns for each price control for household fixed effects and month-of-sample fixed effects. The rightmost columns for each price control for city by month-of-sample fixed effects (*e.g.*, Fresno in January 2010).

The six regressions in Table 1.5 employ two different measures of price: (1) the

[14]The number of heating degrees in a day is equal to the difference between the day's average temperature and 65. Days with average temperatures above 65°F receive zero heating degrees. More formally, we calculate the number of heating degrees for day t with mean temperature $\bar{T}_t$ (in °F) as $\text{HDD}_t = \mathbb{1}\left\{\bar{T}_t < 65\right\} \times \left(65 - \bar{T}_t\right)$. The HDDs variable above is thus $\text{HDDS} = \sum_t \text{HDD}_t/1000$.

household's marginal price during the relevant billing period, and (2) the household's baseline (first-tier) price during the relevant billing period. These two—rather related[15]—measures of price yield considerably different results, differing both quantitatively and qualitatively. The baseline price suggests an elasticity between −0.10 to 0.02, while the marginal price indicates a *positive* demand elasticity between 0.43 and 0.47. The substantial differences across estimates in Table 1.5 suggest at least one of the aforementioned biases are present. Specifically, the fact that the marginal-price based elasticity estimates are positive (implying upward-sloping demand curves), while the baseline-price based estimates are negative, suggests that the *price-is-a-function-of-quantity* flavor of simultaneity is a first-order problem in this context. This interpretation follows from the results due to the fact that baseline prices are not a function of quantity, while marginal prices are a function of quantity.

While the baseline-price based elasticity estimates appear to be reasonable in terms of magnitude, they are still not identified, as they still may suffer from simultaneity bias. Simply adding more observations in the flavor of the *big data* movement does not address this potential endogeneity: column 4 of Table 1.5 does not appear any more plausible than columns 5 or 6, despite adding more than 7 million observations—the same can be said for column 1 *vs.* columns 2 and 3. In addition, the fact that the baseline-price based estimates change sign and magnitude when we move from the 5% CA sample (column 4) to the border-discontinuity motivated sample (columns 5 and 6) provides some evidence that *classical* simultaneity may be present. In this border-discontinuity motivated sample, within-zip code price variation comes from utilities' differentially pricing natural gas over a set of potentially comparable households. However, whether the change in coefficients is due to removing endogenous variation or due to changes in the sample, the existence of simultaneity is fundamentally a statistically untestable issue which stems from the theoretical setup of how market prices originate. Rather than assuming that the sample and/or fixed effects remedy the problem, we instead present a multipart empirical strategy to directly resolve the challenge.

Finally, it is worth noting that the baseline-price based elasticity estimates are well within the range of estimates from the existing literature, as shown in Table 1.1. This outcome warrants some concern, as it suggests that some of these estimates may suffer from endogeneity.

1.4.3 Solutions for identification

Having shown that OLS with fixed effects and extensive data does not cleanly identify the own-price elasticity of demand in this setting, we now discuss several potential routes for identifying the causal effect of price on quantity in our setting. In the end, we opt for an identification strategy that interacts a spatial discontinuity with an instrumental variables approach.

[15]The correlation between marginal price and baseline price is approximately 0.79; see Table A.1 for bivariate correlations of prices measures.

Discontinuities

A common route toward identification in applied microeconomics involves finding relatively small geographic units that receive different prices within the same time period. The assumption is that observable and unobservable characteristics and, more importantly, households' price responsiveness do not differ across the border, yet they are exposed to different prices changes allowing for econometric identification. Arbitrary administrative boundaries that determine policies' catchment areas provide a popular tool in this context, *e.g.*, Dell 2010; Chen et al. 2013; Ito 2014. In our context of natural gas in California, the boundary between PG&E and SoCalGas offers potentially arbitrary within-city (and within-zip code) variation in prices during a month. Specifically, the boundary between PG&E's and SoCalGas's natural gas service areas bisects eleven cities—in three clusters—in southern California: Arvin, Bakersfield, Fellows, Fresno, Del Ray, Fowler, Paso Robles, Selma, Taft, Tehachapi, and Templeton. The left panel of Figure 1.5 displays the two utilities' service areas throughout California (for zip codes sufficiently covered in the datasets). The right panel of Figure 1.5 zooms in on the eleven cities (39 zip codes) that PG&E and SoCalGas both serve. Within these eleven cities, PG&E serves all 39 zip codes, while SoCalGas serves 18 of the zip codes.

This identification strategy rests upon the assumption that households on one side of the utilities' border provide a valid control group for households on the other side of the border. Because the boundary mainly represents the extent of each utilities' underground distribution network and is unlikely to enter into households' preferences, the exogeneity of the boundary to household characteristics should be valid (Ito 2014). The main threat to this identification strategy is that utilities' networks correlate with geographic or neighborhood characteristics over which individuals have preferences. However, we use household fixed effects, which absorb mean differences across households. Thus, for the border discontinuity to be invalid, households would have to sort in a way consistent with their elasticities, and the utilities' price series would have to differ significantly in their variances. Because the data contain considerable variation in prices for both utilities and the panel contains approximately six years of monthly bills, this sort of sorting bias seems unlikely. Figure 1.6b suggests the generating distributions for the utilities' prices are quite similar (the standard deviations of the price series are 0.0940 and 0.1053 for PG&E and SoCalGas, respectively). In addition, Table 1.4 provides some limited evidence[16] of balance across the utility border, comparing PG&E and SoCalGas households within season (summer or winter) and within income group. Within a season-income group, the utilities' customers appear to consume similar volumes of natural gas, receive similar numbers of days per bill, receive similar allowances on the first tier, and face similar numbers of heating degree days. SoCalGas customers tend to receive slightly lower bills, but the difference is less than half of one standard deviation of total bill amount.

Ito 2014 employs a similar strategy within the context of electricity consumption. However, there is at least one significant difference between the electricity

[16] Our data on households is restricted to information from natural gas bills.

and natural gas contexts which prevents us from entirely adopting Ito's identification strategy: discontinuities within electricity utilities' seven-tier pricing regime. By law, the electricity utilities in Ito's study are not allowed to move the price of the first two tiers—they must recover changes in their costs by moving tiers three through seven. In addition, electricity utilities in California do not generally change consumer's prices each month—and prices do not change across all utilities at the same time. Thus, marginal prices in Ito's setting move differently depending upon a household's tier and utility. Ito argues that the residual variation—combining the spatial discontinuity with this pricing discontinuity and spatiotemporal fixed effects—is plausibly exogenous from demand shocks. Because natural gas (in California) has only two tiers and because the absolute difference between the two tiers has relatively low variation, we are unable to take advantage of price-tier based discontinuities. Therefore, in addition to this utility-border-based discontinuity, we adopt an additional strategy to overcome endogeneity.

Instrumental variables

The second element in our empirical strategy for identifying the price elasticity of demand for natural gas involves a traditional solution to simultaneity: supply-shifting instruments. In this context, the ideal supply-shifting instrument is (1) strongly correlated with the prices that the natural gas utilities charge their customers (the *first stage*), and (2) uncorrelated with residual shocks affecting consumers' demand (Angrist and Pischke 2009). In this paper, our instrument is the Henry Hub spot price for natural gas.

Henry Hub spot price Specifically, we instrument the prices that consumers face (*e.g.*, marginal price, average price, baseline price) with the average spot price at Louisiana's Henry Hub in the week preceding the change in prices. We also interact the Henry Hub spot price with *utility* to allow the utilities to differentially pass through price changes. The Henry Hub spot price represents the nationally prevailing price for short-term natural gas contracts (the hub sits at the intersection of 13 intrastate and interstate pipelines) (U.S. Energy Information Administration 2016a). This instrument mechanically satisfies the requirement of having a strong first stage, as both utilities base their prices, in part, on market prices for natural gas in the period preceding their rate changes—the utilities buy natural gas on the spot market, and the California Public Utilities Commission regulates how the utilities fold their costs into the price regimes that customers face on a monthly basis.

The exclusion restriction for this spot-price based instrument is less obvious, but several factors suggest the exclusion restriction is plausibly valid. First, California's entire residential natural gas demand represents *at most* three percent of national natural gas consumption—limiting the individual utilities' ability to set/influence spot prices and the Henry Hub. Second, we interact the spot price instrument with utility. This interaction, conditional on city by month-of-sample fixed effects, implies that the identifying variation in our instruments comes from the difference in how the two utilities' incorporate monthly spot-price shocks into their pricing regimes. Third, because the utilities must obtain approval for price changes before the new price regime begins, the spot price is temporally disconnected from the billing period. In other word, the utilities' costs (and approved prices) are based upon spot prices that precede the billing period by several weeks. Thus, shocks that affect the Henry Hub spot price are distinct in time from shocks that affect natural gas demand—our fixed effects will absorb any of these shocks, so long as they do not differ across the utilities' border within a month.

In addition, we show that the most salient lag of price is likely the second lag of price, further disconnecting contemporaneous local demand shocks from market-level supply shocks two months prior.[17] We also control for the number of heating degree days (HDDs) in the household's zip code during the households' billing period. Because residential consumers primarily use natural gas in heating applications, controlling for HDDs further reduces the opportunity for local demand shock to affect national prices. One final exclusion-restriction concern is that price variation in the spot market for natural gas may affect both residential natural-gas prices *and* residential electricity prices. In this scenario, we would not be able to separate the effect of an electricity price shock from a natural-gas price shock. However, Figure 1.8 suggests that (1) residential natural gas and electricity prices are uncorrelated in both levels and differences (across the utilities' border within a month of sample), and (2) variation in the residential price for electricity is uncorrelated with variation in the Henry Hub spot price of natural gas.[18] Therefore, we argue that the exclusion restriction is plausibly valid for our spot-price instrument.

Our identification strategy thus interacts the spatial discontinuity between PG&E's and SoCalGas's service areas with the Henry Hub spot price. Specifically, the identifying variation stems from the two utilities' divergent pass-through of the spot market price—differentially projecting variation in the the natural gas spot market across a tenably arbitrary border between the two utilities.

By employing a two-part identification strategy that interacts a spatial discontinuity with a price instrument, we avoid weaknesses inherent in either individual identification strategy. For instance, simply instrumenting residential prices with the spot-market price may not entirely purge the endogenous, *bad* variation from residential prices, as variation in the spot market likely results from both supply and demand shocks. Our identification strategy instead takes variation from the spot market and projects it across the utilities' border, treating neighboring households with prices that differ only due to which utility provides natural gas. Additionally, our identification strategy also allows repeated "treatments" across the discontinuity, as the utilities change residential natural gas prices each month. This repetition of treatment both increases power and diminishes concerns regarding sorting, as both sides of the border will be "treated" over time. Thus, we contend this two-part identification strategy is well-suited for the challenges to identification in this setting.

[17]See Tables 1.7, A.3–A.7 for the second-stage results comparing consumers' responses to various lags in price.

[18]This observation also draws upon Figure 1.6b.

Spot price IV, first stage Panel A of Table 1.6 provides the first-stage estimates for the two-stage least squares (2SLS) equations

$$\log(p_{i,t}) = \pi_{1a} p_{i,t}^{\text{spot}} + \pi_{1b} p_{i,t}^{\text{spot}} \times \text{SCG}_i + \pi_2 \text{HDD}_{i,t}^{\text{bill}} + \text{HH}_i + \text{City}_{i,t} + u_{i,t} \tag{1.2}$$

$$\log(q_{i,t}) = \eta_1 \widehat{\log(p_{i,t})} + \eta_2 \text{HDD}_{i,t}^{\text{bill}} + \text{HH}_i + \text{City}_{i,t} + v_{i,t} \tag{1.3}$$

where HH_i is a household fixed effect, $\text{City}_{i,t}$ is a city by month-of-sample fixed effect, SCG_i is an indicator for whether the household's retail utility is SoCalGas, and $\text{HDD}_{i,t}^{\text{bill}}$ is the number of heating degree days for household i during its billing period that began in month t.

The spot-price by utility term is not collinear with the city by month-of-sample fixed effects because households' bills do not perfectly align with calendar months: consequently, the bills span multiple price regimes (see Figure 1.7). In a bill that spans two calendar months, the household receives a weighted average of the two months' price regimes. To instrument these weighted averages of prices, we calculate corresponding weighted averages of the relevant spot prices by weighting the months' relevant spot prices by their temporal share of the bill.[19]

Figure 1.6b provides visual evidence of the first stage—illustrating (1) the link between the two utilities' prices and the Henry Hub spot price and (2) the utilities' differential responses to the spot price. Throughout the rest of the paper, we define the Henry Hub spot price as the average spot price for natural gas at the Henry Hub during the seven days preceding the utility's change in pricing.

Panel A of Table 1.6 displays the first-stage results for equation 1.2 using five different prices that may be relevant to households: marginal price, average price, average marginal price, baseline price, and simulated marginal price[20] (using the log of each price). Each price is the second lag of the contemporaneous price.[21] Table 1.7 and Tables A.3–A.7 compare consumers' varying responses to different lags of price.

Both Figure 1.6b and Panel A of Table 1.6 demonstrate that the spot-price based instruments are quite strong: the F statistics testing the joint significance of the instruments range from 369.9 to 1,333.2. This significance is unsurprising, because the utilities purchase gas on the spot market and incorporate these costs directly into their price regimes. The significance of the interaction between spot price and utility (SoCalGas) in the second row of Panel A in Table 1.6 confirms that the utilities differ appreciably in incorporating spot-market costs into their pricing regimes: PG&E's pricing regime appears to be much less responsive to the contemporaneous spot price than that of SoCalGas.[22] Though the city by month-of-sample fixed effect

[19]See the appendix section Calculating bills for further discussion of bills spanning multiple months.

[20]*Simulated marginal price* refers to a simulated instrument for marginal price. We discuss this measure of price in the next section.

[21]The current bill is lag zero, the prior bill contains the first lag of price, and the bill preceding the prior bill contains the second lag of price.

[22]One difference between the utilities' pricing regimes is that PG&E does not have a fixed charge, while SoCalGas does. Thus, PG&E recovers both fixed and volumetric costs through volumetric

should control for most local demand shocks, bills do not perfectly match months. The within-bill HDDs variable $\text{HDD}_{i,t}^{\text{bill}}$ in equation 1.2 controls for any remaining weather-based demand shocks. The results in Table A.10 demonstrate robustness to excluding (odd-numbered columns) or including (even-numbered columns) within-bill heating degree days, which suggests that the instrument is exogenous to local weather shocks, one of the key local-demand drivers in natural gas (Davis and Muehlegger 2010; Levine, Carpenter, and Thapa 2014; Hausman and Kellogg 2015).

While the first stage is quite strong for all specifications, the results in Panel A of Table 1.6 suggest the instrument is strongest—in terms of first-stage significance—for baseline price, followed by average marginal price, average price, marginal price, and finally simulated marginal price. A likely reason for this outcome is that baseline price is the least noisy price: it is the only price that is not a function of the consumer's quantity, and it does not include variation from changes in the size of the step between the two tiers' prices. By these terms, (simulated) marginal price is the noisiest, which is consistent with marginal price having the smallest first-stage F statistic of the five prices.[23]

Instrumented prices and simulated instruments

In the preceding sections, we discussed how we interact a spot-price-pass-through based instrument with a spatial discontinuity in utilities' service areas to overcome bias stemming from the classic form of simultaneity—*i.e.*, quantity and price (our dependent and independent variables) result from a simultaneously determined equilibrium. We now discuss the aspect of our identification strategy that deals with the *price-is-a-function-of-quantity* endogeneity present in multi-tiered pricing contexts.[24] We present three separate options for breaking this endogenous link between price and quantity, but, in the end, the options yield very similar results.

Option 1: Instrumented prices One method for breaking the endogenous link between a household's price and its quantity is simply to instrument the household's price with a variable that is aggregated at a unit above household. Consider the IV strategy discussed above: instrumenting a household's price with the Henry Hub spot price interacted with utility. Because this instrument only varies at the level of billing-period by utility, when we regress a household's endogenous price on this instrument (and our set of fixed effects) in the first stage, the variation captured by the predicted prices is only the variation that correlates with the spot price, which is determined weeks, if not months, before the household's consumption decision. Thus, if the spot price provides a valid instrument in the classical simultaneity

charges to its customers.

[23] Although the first-stage estimates in Panel A of Table 1.6 have the flavor of pass-through results, one should keep in mind that equation 1.2 specifies a log-linear form (logged price as the response variable), which does estimate pass-through.

[24] This endogeneity is present in marginal price, average price, and average marginal price—all three prices are functions of the individual household's quantity consumed.

context, it also provides a valid instrument for the second *price-is-a-function-of-quantity* endogeneity.

Option 2: Baseline price In a similar manner, the baseline price provides a valid instrument that breaks the *price-is-a-function-of-quantity* endogeneity. Because a household's baseline price is not a function of its quantity consumed, baseline price does not suffer from the same endogeneity. Baseline price is also strongly predictive of marginal (or average) price.[25] Thus, in application, one could either replace marginal (or average) price with baseline price *or* instrument one of the endogenous prices with baseline price. There is at least one drawback to this approach: baseline price, by definition, fails to capture the higher price that a household faces when the household exceeds its total monthly allowance.

Option 3: Simulated instrument Simulated instruments[26] provide a third option for breaking the *price-is-a-function-of-quantity* flavor of endogeneity. The simulated-instrument approach follows a methodology suggested in Ito 2014. Specifically, this approach creates an instrument (or proxy) for marginal (or average) price by plugging a lagged level of consumption into the current price regime, *i.e.*,

$$z_{i,t} = p_{i,t}\left(q_{t-k}\right) \tag{1.4}$$

The main idea for this instrument is using a household's consumption history to predict whether a household will face the baseline or excess price in the current period. As with any instrument, we want to accomplish this prediction in a way that is strongly predictive of the true outcome (the first stage) and that is uncorrelated with any recent shocks to the household (the exclusion restriction) (Angrist and Pischke 2009). For these reasons, we modify equation 1.4 slightly. First, we use the households' lagged consumption levels (from lagged bills 10 through 14 months prior) to calculate the share of lagged periods that exceed this billing period's baseline allowance, *i.e.*,

$$s_{i,t} = \frac{1}{5}\sum_{k=10}^{14} \mathbb{1}\{q_{i,t-k} > \bar{A}_{i,t}\} \tag{1.5}$$

where $\bar{A}_{i,t}$ is household i's baseline allowance in time (bill) t. We then calculate the *simulated instrument* for marginal price, $z_{i,t}$, as

$$z_{i,t} = \mathbb{1}\{s_{i,t} \leq 0.5\} \times p_{i,t}^{\text{base}} + \mathbb{1}\{s_{i,t} > 0.5\} \times p_{i,t}^{\text{excess}} \tag{1.6}$$

Summarizing equations 1.5 and 1.6: this simulated instrument for marginal price predicts that a household will exceed its allowance when the majority of the household's past bills (using lagged months 10 through 14) exceed the current bill's

[25] The correlation between baseline price and marginal price is approximately 0.79; the correlation between baseline price and average price is approximately 0.94. See Table A.1 for all bivariate correlations between our five measures of price.

[26] Also called *policy-induced instruments*.

allowance.[27]

Table A.2 provides evidence that this *simulated-instrument* approach significantly predicts households' marginal prices. Specifically, Table A.2 provides the estimate and standard error for β in the equation

$$p_{i,t}^{\text{mrg}} = \beta p_{i,t}^{\text{sim}} + \text{HH}_i + \text{City}_{i,t} + w_{i,t} \tag{1.7}$$

where $p_{i,t}^{\text{mrg}}$ is household i's marginal price in time t and $p_{i,t}^{\text{sim}}$ is our simulated instrument for household i's marginal price in time t $\left(\text{i.e., } p_{i,t}^{\text{mrg}}\right)$. The estimates for β in Table A.2 confirm the strong "first stage" for this simulated instrument. Marginal price and simulated marginal price are strongly and significantly correlated—both t statistics are approximately 148. The two columns in Table A.2 also provide evidence of the robustness of the simulated instrument to the choice of lags: the estimates using lags 10–14 or 11–13 are virtually indistinguishable. In addition, the bottom row of Table A.1 demonstrates that this simulated instrument is strongly correlated with marginal price ($r \approx 0.85$) in addition to the other four measures of price.

Column 5 of Table 1.6 (Panel A) provides the first-stage results consistent with equation 1.2 but with the simulated instrument of marginal price substituted (proxying) for actual marginal price (and still instrumenting with spot price interacted with utility across the utilities' border).[28] The first stage is again quite strong in this specification, and the results are qualitatively similar to the results in columns 1–4 of Table 1.6, Panel A. Henceforth we will refer to the simulated instrument for marginal price as *simulated marginal price*.

All subsequent results apply our two-part identification strategy which exploits the utilities' differential pass-through of spot-market prices to obtain exogenous variation in residential natural gas prices across the border between the two utilities' service areas. To incorporate the three competing options discussed immediately above, we provide results consistent each the strategies: instrumenting with spot price interacted with utility, proxying with baseline price, and employing simulated marginal price (the simulated instrument/proxy for marginal price). We now turn to our main results.

1.5 Results

In this section, we discuss the estimated price elasticities, using the empirical strategies extensively discussed in the preceding section. After presenting the main results

[27]This simulated instrument is robust to the choice of months 10 through 14. The goal is to keep the instrument in the same season as the current bill (maintaining a strong first stage), while allowing some temporal distance between the lags and the current period (the exclusion restriction: preventing short- and medium-run shocks from affecting both periods).

[28]It is worth noting that, in this paper, any result using the simulated instrument will have fewer observations than other results, as the simulated instrument is greedier for data—for an observation to remain in the dataset, its 14$^{\text{th}}$ lag must also be in the dataset. Our other price measures are not as greedy.

for the *pooled* elasticity (no heterogeneity), we examine whether households' price responses (*i.e.*, elasticities) vary by season and/or by income.

1.5.1 Pooled price elasticity of demand for natural gas

Panel B of Table 1.6 displays the elasticity results from the second-stage regression specified in equation 1.3. These results instrument log price with the Henry Hub spot price (interacted with utility), exploit the spatial discontinuity, and use the log of daily average consumption (in therms) as the outcome. The five columns each estimate the elasticity using the log of a different type of price: marginal price, average price, average marginal price, baseline price, and simulated marginal price. As discussed above, each price is the second lag of price, as opposed to the contemporaneous price. The estimates for the price elasticity of demand range from −0.17 (simulated marginal price) to −0.23 (average price).

Panel B of Table 1.6 indicates that the estimated elasticity is fairly robust to the type of price. Table A.11 demonstrates that the estimated elasticity is also robust to the inclusion/exclusion of heating degree days and to the levels of fixed effects—ranging from city by month-of-sample fixed effects to zip-code by week-of-sample fixed effects (while still including household fixed effects). The robustness to type of price also demonstrates robustness to how we control for the *price-is-a-function-of-quantity* endogeneity discussed above. Tables A.11–A.15 demonstrate the robustness of the estimated elasticity to excluding within-bill heating degree days and varying the spatiotemporal fixed effects. Finally, Table A.16 contains marginal-price based elasticity estimates as we incrementally extend the study-area. Beginning with the study area (*Common Zips*), we add the zip codes neighboring (bordering) the study area (*Neighbors 1*); we then add the neighbors of the neighbors (*Neighbors 2*); last, we add a third band of neighbors (*Neighbors 3*). Figure A.4 illustrates these groups of neighboring zip codes. The estimated elasticity from the first group of neighbors (−0.19 (0.05) in column 2 of Table A.16) is quite close to the elasticity previously discussed (−0.21, (0.07) in column 1); the elasticities that include the second and third peripheral neighbors diminish in magnitude (−0.12 and −0.09) but still differ significantly from zero.

Compared to their OLS-based counterparts in Table 1.5, the marginal-price based 2SLS estimates for the elasticity of demand now have opposite—and theoretically correct—signs. The magnitudes of the 2SLS estimates of the elasticity (approximately −0.20) are theoretically reasonable and within the range of previous findings. Furthermore, these estimates are plausibly identified and utilize consumers' actual prices.

As discussed above, the results discussed up to this point—*e.g.*, the results in Table 1.6—estimate the price elasticity of demand for residential natural gas using the second lag of the various prices. In order for a household to know the prices of its contemporaneous bill, the household would need to closely follow the approved advice letters published online by the utility or the California Public Utilities Commission. Otherwise, the household will learn about prices from past bills—hence the use of lagged prices. Because a household will not receive the bill for the previous billing period for several days into its current billing period—and because the household may not view the previous bill until it pays the bill (or its credit card bill, if the household uses automatic bill payment) weeks later—the household may not know the prices from its immediately previous bill until the current period is nearly over. For these reasons, it is plausible that the second lag of price is the most salient price to many households. Figure 1.7 illustrates an example of the timing for bill delivery, bill payment, and the relevant lags of prices.

Table 1.7 replicates the second-stage results for marginal price and average price but varies the lag/lead of price: beginning with the first lead of price, followed by contemporaneous price, the first lag of price, and finally the second lag of price. Tables A.3–A.7 provide further detail, varying the lead/lag of each of the five prices—ranging from the first lead of price to the third lag of price. Across the five types of measures of price, none of the first leads of price, contemporaneous prices, or first lags of price differ significantly from zero. For each type of price, both the second and third lags of price differ significantly from zero. For each price, the second-lag based elasticity slightly exceeds the third-lag based elasticity in magnitude, but the difference does not exceed the standard error. These results are consistent with households responding to two-to-three lags of price—as opposed to contemporaneous price—suggesting some degree of inattention by the household to the true price, akin to previous work on inattention and salience, *e.g.*, Chetty, Looney, and Kroft 2009; Sallee 2013; Allcott and Taubinsky 2015.

1.5.2 Heterogeneity

We now examine evidence of heterogeneity in the price elasticity of demand for natural gas. The institutional setting of this paper motivates two relevant dimensions of heterogeneity—income level and season—as the CPUC and utilities already apply different price regimes to households depending upon (1) the season of year (summer *vs.* winter) and (2) the household's income level (specifically, CARE status). If heterogeneity exists, then the regressions in the preceding section *pool* across the heterogeneous effects. This pooled parameter estimate may still be relevant for policy applications—particularly for policies that cannot differentiate between seasons and/or income groups. However, because OLS weights heterogeneous treatment effects by their shares of the residual variation in the variable of interest—which is itself a function of (1) the numbers of observations in the heterogeneous groups and (2) the (residual) within-group variance in the variable of interest (Solon, Haider, and Wooldridge 2015)—one might wonder whether the pooled estimator always provides a policy-relevant estimate. In addition, in the presence of heterogeneous elasticities, policymakers can increase efficiency by integrating these (known) heterogeneities (Ramsey 1927; Boiteux 1971; Davis and Muehlegger 2010).

For income-based heterogeneity, we use a household's CARE status as a proxy for its income level.[29] As discussed above, households qualify for CARE by either

[29]Because we do not have identifying variation in income level (or season), the heterogeneities that we estimate should be taken as descriptive for the given group, rather than causal effects of

meeting low-income qualifications or by receiving benefits from one of several state or federal assistance programs (*e.g.*, Medi-Cal or the National School Lunch Program) (Southern California Gas Company 2016). For seasonal heterogeneity, we split the calendar into winter months (October through March) and summer months (April through September).[30]

Income heterogeneity

To examine income-based heterogeneity in the price elasticity of demand for natural gas, we estimate the two-stage least squares equations 1.2 and 1.3 separately for CARE households and non-CARE households. Columns (3) and (4) of Table 1.8 supply the second-stage results from these regressions, providing estimates of the elasticity of demand by income level (CARE status).

The results in columns (3) and (4) of Table 1.8 suggest that the elasticity results in the previous section may in fact pool across heterogenous elasticities; we estimate that the price elasticity for CARE (lower-income) households is approximately twice that of non-CARE (higher-income) households. Specifically, using the marginal price, we estimate an elasticity of approximately –0.24 (0.080) for CARE (lower income) households and –0.14 (0.068) for non-CARE households. The "pooled" estimate corresponding to these results is –0.21 (0.071) (column (1) of Panel B in Table 1.6)—slightly higher than the midpoint between the CARE estimate and the non-CARE estimate.

Seasonal heterogeneity

To estimate seasonal heterogeneity in the price elasticity of demand for residential natural gas, we estimate the two-stage least squares equations 1.2 and 1.3 separately for winter months and for summer months. Columns (1) and (2) of Table 1.8 supply the second-stage results from these regressions, providing estimates of the elasticity of demand by season.

The results in columns (1) and (2) of Table 1.8 indicate a stark and significant difference between price elasticities in summer and winter months. Using marginal price, we estimate that the price elasticity of demand for natural gas in summer months is approximately 0.052 (0.029), which marginally differs from zero at the 10 percent level. The estimated elasticity for winter months is approximately –0.38 (0.14) and differs significantly from zero at the 1 percent level. The comparable "pooled" elasticity estimate corresponding to these results is approximately –0.21 (0.071) (column (1) of Panel B in Table 1.6). These results provide strong evidence

income level or season. In other words, while we estimate heterogeneous elasticities with respect to income level, this heterogeneity may have nothing to do with income and could instead result from some other (omitted) variable that correlates with income/CARE status, *e.g.*, the age and size of the physical home. However, identification of the sources of heterogeneity is not the goal of this paper; we aim to identify the elasticity of demand and demonstrate dimensions of heterogeneity. We leave it for future papers to identify the sources of these heterogeneities.

[30] This definition reflects southern California's two seasons: warm and slightly less warm.

that households' consumptive and price-response behaviors vary considerably by season—the winter-based elasticity is nearly twice the "pooled" elasticity.[31][32]

Income-by-season heterogeneity

Having shown potential heterogeneity across income groups (CARE status) and season, we now examine the evidence that income groups' heterogeneity varies by season by interacting the two heterogeneity dimensions discussion above (income and season).

To estimate seasonal-by-income heterogeneity in the own-price elasticity of demand for residential natural gas, we estimate the two-stage least squares equations 1.2 and 1.3 separately for the four potentially heterogeneous subgroups: CARE households in the summer, non-CARE households in the summer, CARE households in the winter, and non-CARE households in the winter. Table 1.9 displays the second-stage results from these regressions, providing estimates of the elasticity of demand by season and CARE status.

The results in Table 1.9 are consistent with heterogeneous elasticities that depend upon the interaction between household income (CARE status) and season. In other words, the difference between a household's winter and summer price elasticities varies by the household's income level. Specifically, the results in Table 1.9 indicate that both income groups are essentially inelastic to prices in summer months; we estimate a "summertime" price elasticity of 0.046 (0.035) for CARE households and 0.074 (0.032) for non-CARE households. Both elasticities are positive, but only one is significantly different from zero and small. In winter months, both sets of consumers are significantly and substantially more responsive to price, but CARE households are particularly price responsive. We estimate the "wintertime" price elasticity of demand for natural gas is −0.523 (0.142) for CARE households and −0.317 (0.150) for non-CARE households. Again, the pooled elasticity corresponding to these results is approximately −0.21 (0.071) (column (1) in Table 1.6), which is a bit lower than the average of these four elasticities. Overall, Table 1.9 demonstrates the potential for substantial and important heterogeneity underlying commonly estimated pooled elasticities.

1.6 Discussion and conclusion

This paper combines millions of household natural gas bills with a multi-part identification strategy to provide the first micro-data based causal estimates of the own-price elasticity of demand for residential natural gas. Utilizing cross-border price variation between California's two largest natural gas utilities— resulting from the

[31] Table A.8 reproduces these heterogeneity results using average price—rather than marginal price—with very similar results.

[32] Because the current/relevant natural gas institutions divide the year into *winter* and *summer*—and because gas is primarily used for heating—we believe this summer/winter split is the most policy-relevant temporal disaggregation of the price elasticity of residential natural gas. We do not further disaggregate in time.

utilities differential pass-through of spot-market price variation—we isolate plausibly exogenous variation in residential natural gas prices. We estimate an elasticity of −0.21 [−0.35, −0.07]. This estimate is robust to specification choices that include within-bill weather, several price instruments, and the definition/type of price. The point estimates for the own-price elasticity range from −0.23 to −0.17 across five measures of price. Given the robustness of these findings, this paper provides tight bounds on a policy-relevant parameter key to applications ranging from estimating the welfare benefits of fracking (Hausman and Kellogg 2015) to analyzing the regressivity of two-part tariffs (Borenstein and Davis 2012). Because households respond significantly to price changes two to four months prior to the period of consumption—and following Ito 2014—we interpret these estimated elasticities as fairly *medium-run* elasticities.[33]

As a second important finding, we estimate that the own-price elasticity of demand varies significantly across seasons and customer types. We show that households on a popular low-income program, which subsidizes households' natural gas and electricity, appear to be twice as elastic in their response to price as households who are not part of the program. We also show that the price elasticity varies greatly across seasons. If we average across types of households, the summer price elasticity is close to, and only marginally different from, zero. The winter price elasticity is −0.38. This heterogeneity suggests that households are much more price sensitive during their high-consumption months—the winter. These high-consumption winter months also correspond to the time of year in which consumers use natural gas in its most salient form: heating. When we break down the price elasticity across users and seasons, we show that subsidized consumers display the largest price sensitivity during the winter (−0.52). Neither type of customer displays a substantial price response in the summer. These results suggest that, if suppliers want to pass through costs to (or tax) consumers, summertime may be best—both for efficiency and for progressivity. This point hinges critically on the assumption that external costs from natural gas combustion are properly priced. For global pollutants, this is the case in California because the natural gas sector is part of California's cap and trade system.

Figure 1.9 illustrates the seasonal heterogeneity point with simple linear demand that is quite inelastic in the summer and moderately elastic in the winter—consistent with our results. The top row of Figure 1.9 demonstrates that, in this scenario, deadweight loss is substantially larger in the winter than in the summer. The bottom row simply doubles the summer tax and halves the winter tax, resulting in a minuscule increase in deadweight loss for the summer and a substantial reduction in deadweight loss in the winter—implying a considerable overall reduction in deadweight loss.[34] Again, it is worth noting that this example also assumes (1) a first-best world (no unpriced costs to consumption) and (2) the goal of the policymaker is reducing deadweight loss conditional on some level of tax/cost re-

[33]Ito also notes that the medium-run elasticity is often the most policy-relevant elasticity.

[34]This toy example is meant to illustrate an idea. The most efficient seasonal tax adjustment—conditional on a level of tax recovery—would likely not imply exactly doubling taxes in the summer and halving taxes in the winter.

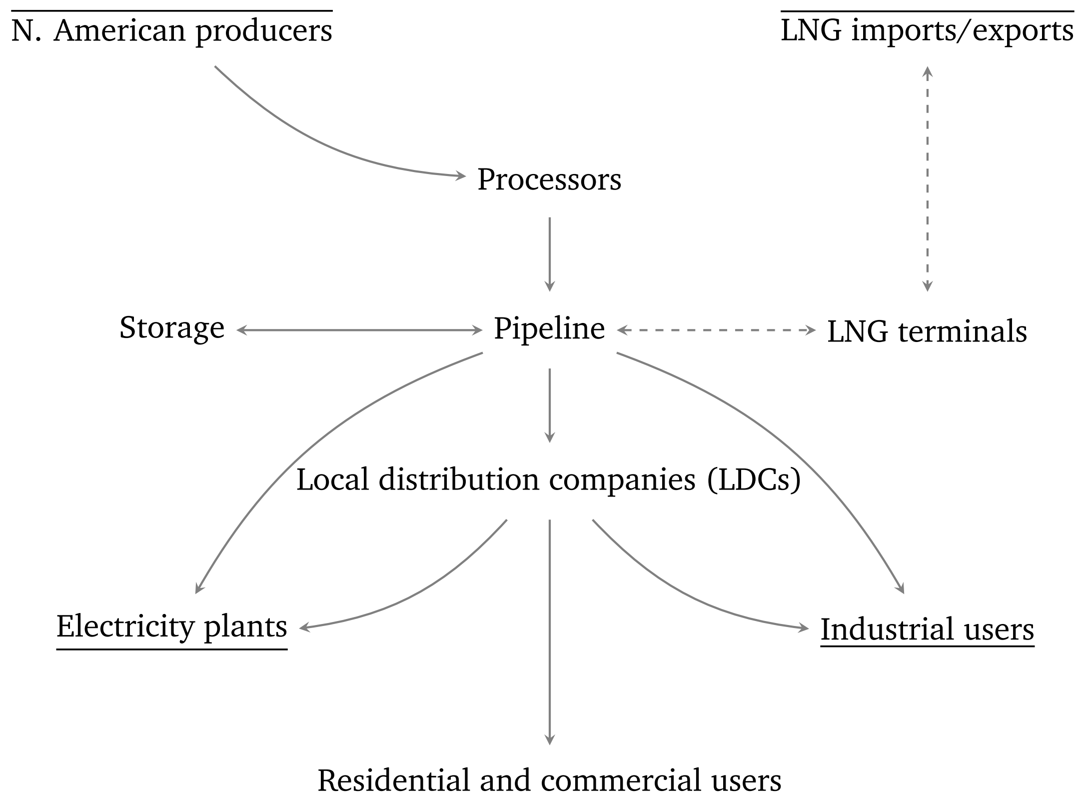
covery. If, for instance, natural-gas consumption includes an unpriced social cost, then increasing summer taxes and reducing winter taxes could potentially further reduce market efficiency by exacerbating the unpriced costs. Similarly, if the policy-maker wishes to use the tax to reduce consumption, then our results suggest that imposing a per-unit tax in the winter is much more efficient than the same tax in the summer.[35] However, our season-by-income results imply that the poor would bear the largest deadweight loss for such a tax.

The discussion above suggests a dimension for tax and cost-recovery efficiency—season of year—that we have not seen recommended in the literature or applied in practice. This fairly simple idea raises a wider question for future work: Along which other dimensions of consumer heterogeneity might we optimize current tax and cost-recovery policies? If policy is to take seasonal heterogeneity—or any other heterogeneity—into account, future work should decompose traditionally *pooled* elasticities and policy responses. Such work will provide policymakers with important parameters to improve market efficiency and enhance policy progressivity.

[35] In terms of units of abatement per dollar of tax levied.

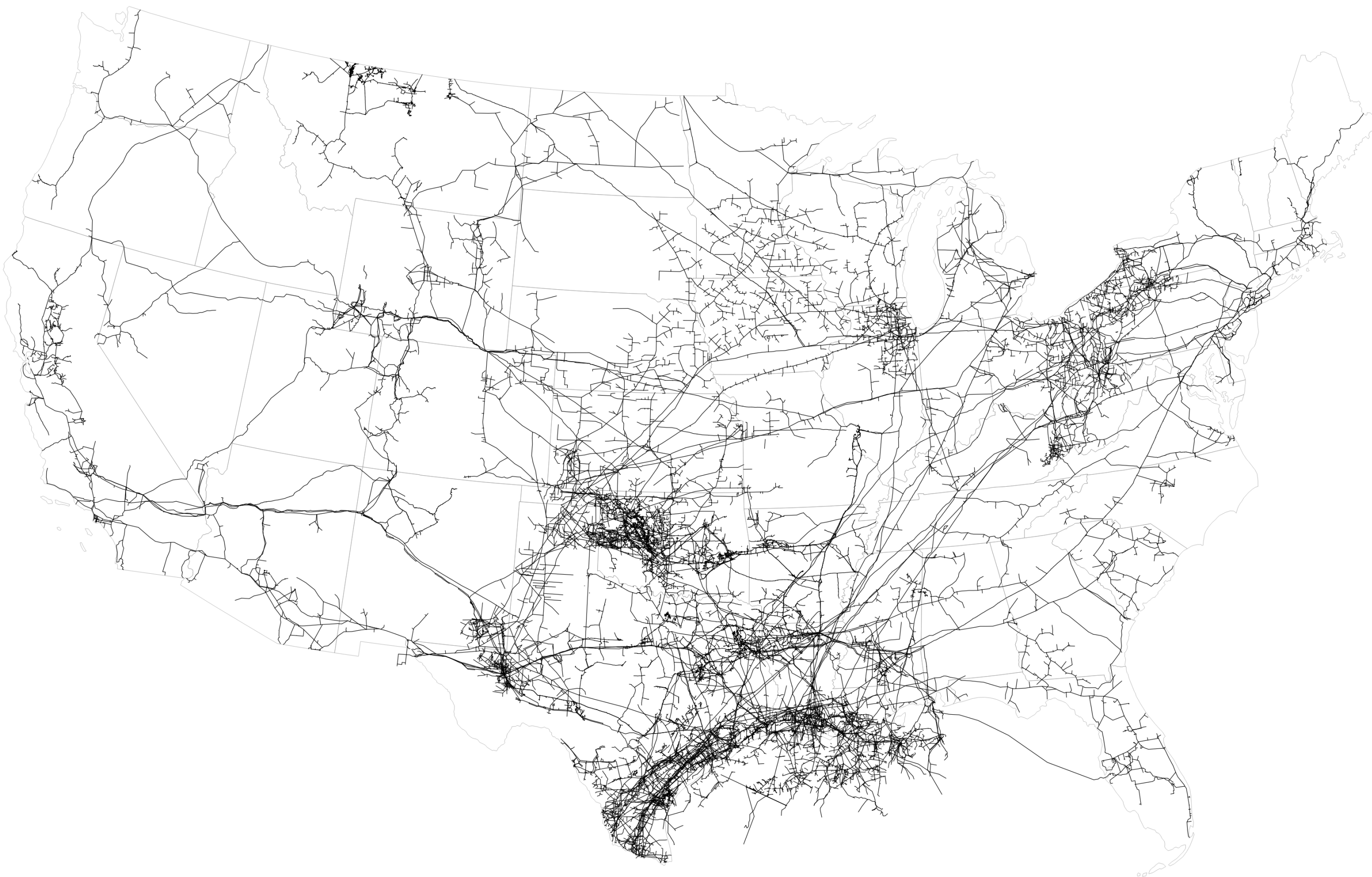
1.7 Figures

Figure 1.1: **U.S. natural gas institutional organization**



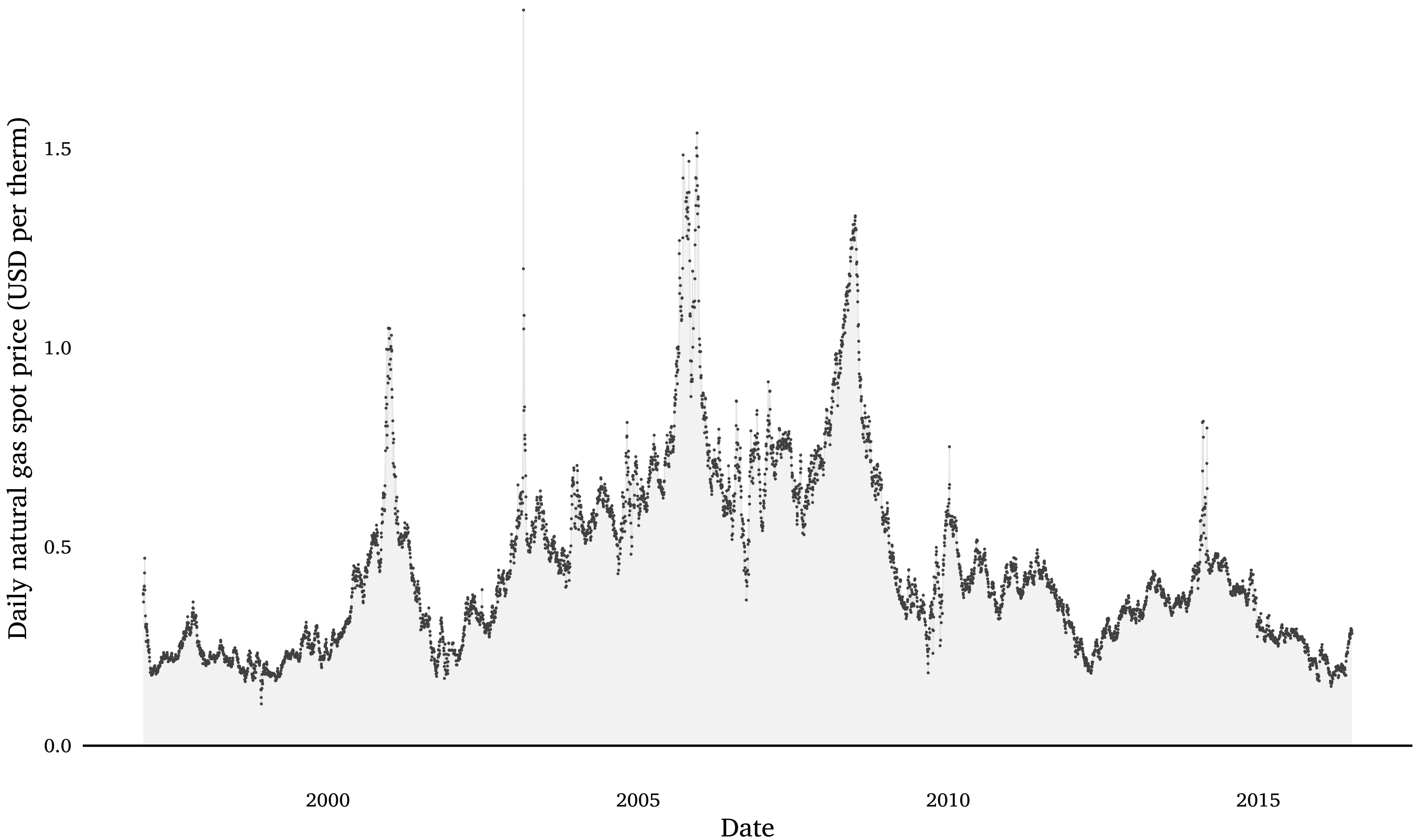
Notes: Overbars represent points of entry into the U.S. natural gas market; underbars represent end points in the market; all other labels represent intermediaries. Arrow directions correspond to the direction of the flow of natural gas. The acronym *LNG* abbreviates *liquid natural gas*. This figure is based on Levine, Carpenter, and Thapa 2014 with modification following Brown and Yücel 1993.

Figure 1.2: **U.S. natural gas pipeline network**



Notes: This map depicts the intra- and inter-state natural gas pipeline network for the (continental) United States (in black) overlayed on a map of the (continental) U.S. (light gray). *Source:* U.S. Energy Information Administration

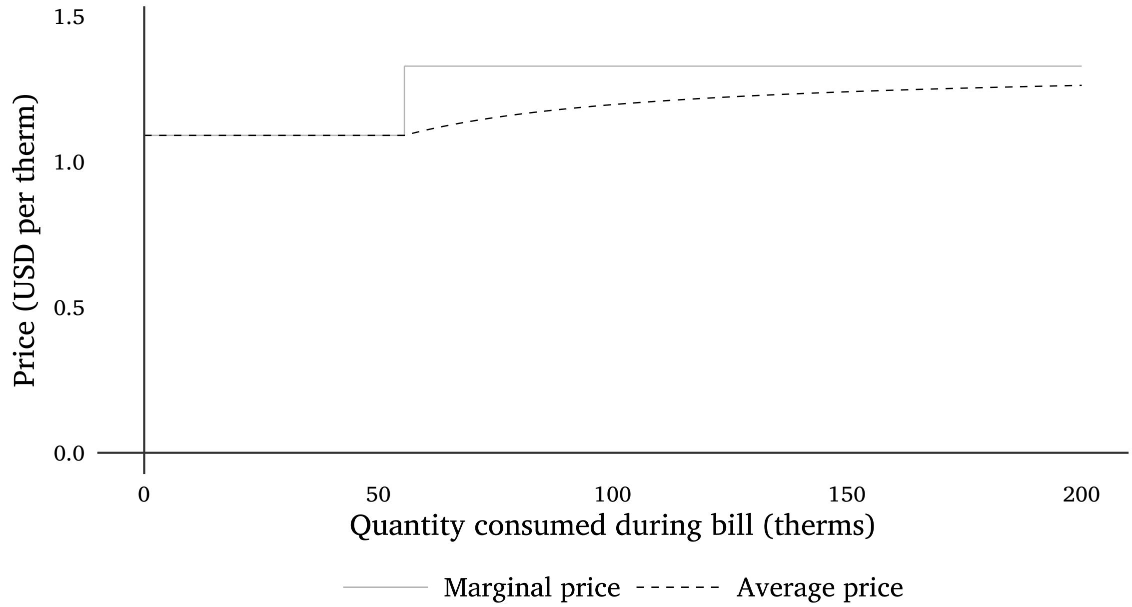
Figure 1.3: **Henry Hub natural gas spot price:** Daily, 1997–2016



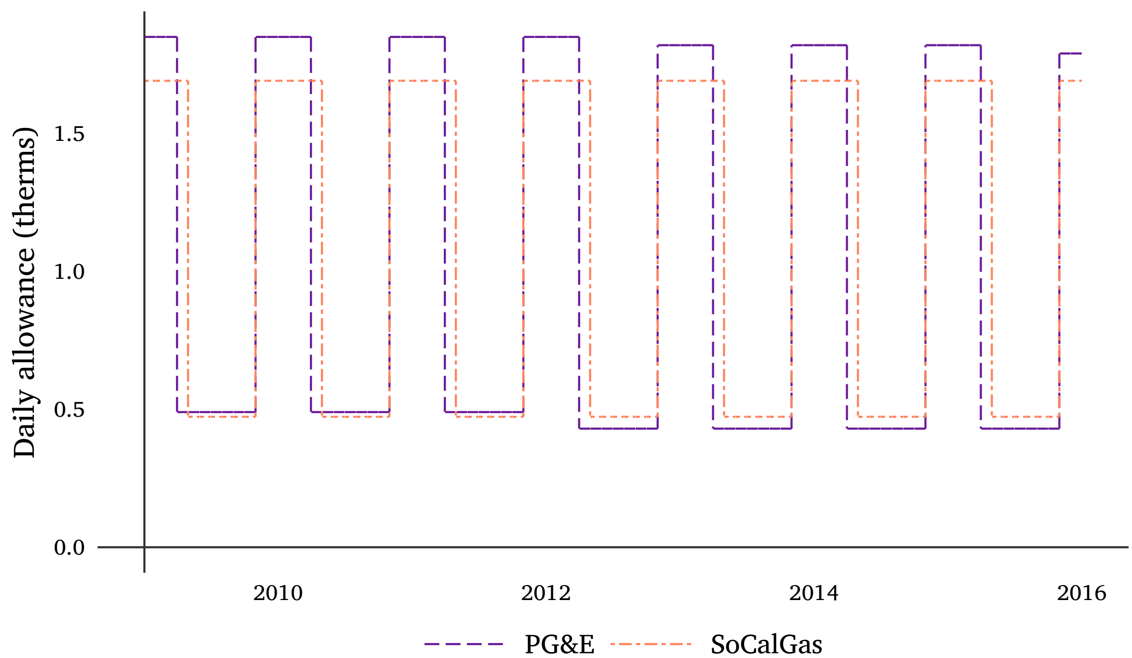
Source: U.S. Energy Information Administration

Figure 1.4: **Households' allowances and prices**

(a) **Allowance and marginal *vs.* average price example:** PG&E, January 2009, climate zone R

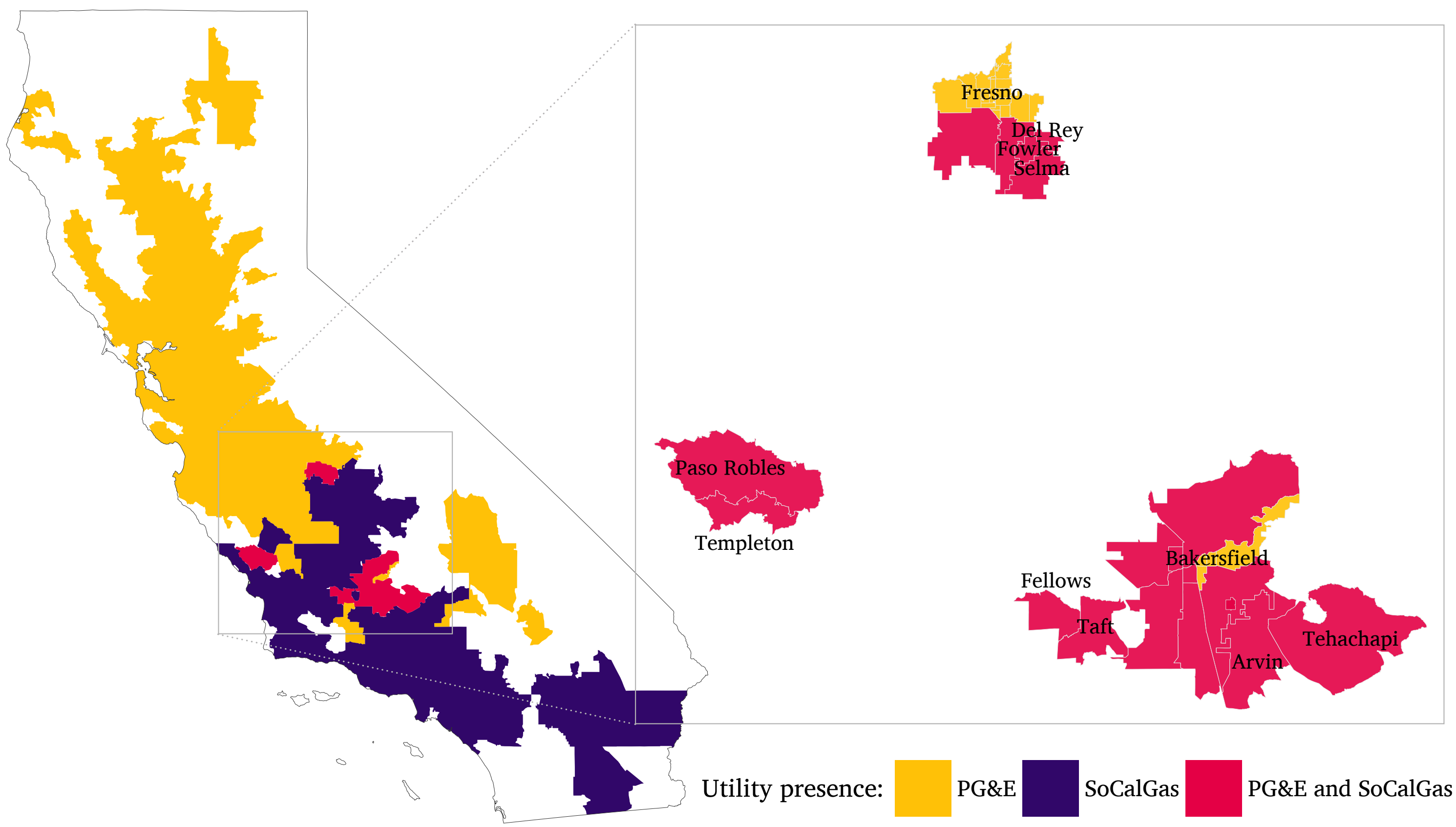


(b) **Tier-one daily allowances over time:** PG&E (zone R) and SoCalGas (zone 1), 2009–2015



Notes: Households receive daily allowances for baseline (first-tier) consumption as a function of location and season (*e.g.*, climate zone *R*, January 2009). The household pays the second-tier price on all units that exceed its allowance—comparing total consumption (during the billing period) to total allowance (daily allowance summed across the bills' days).

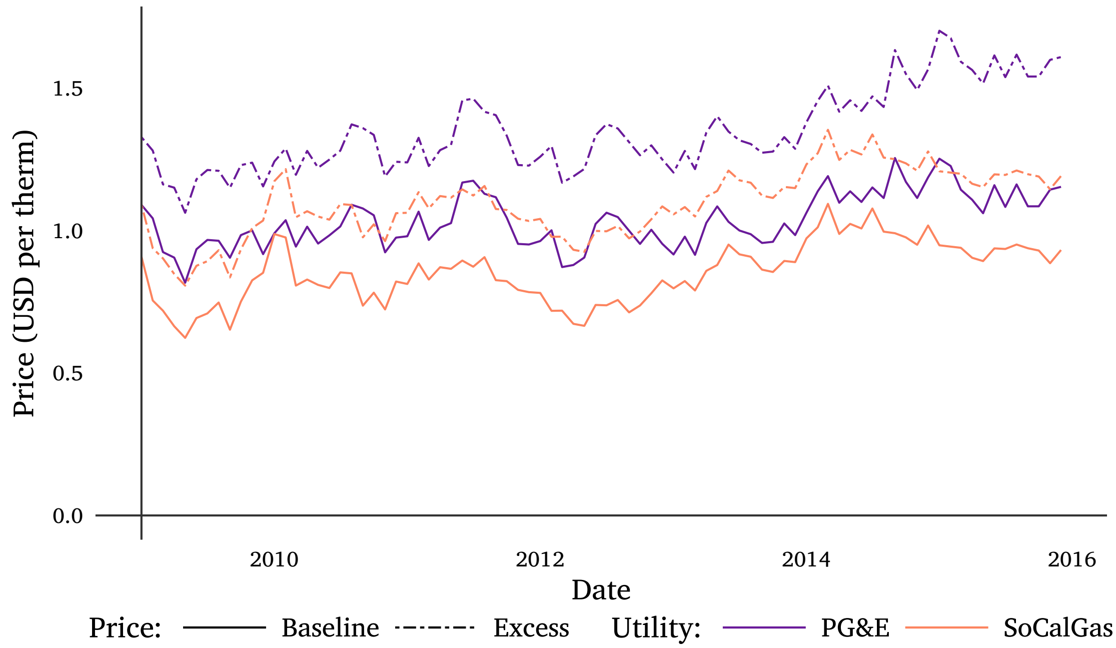
Figure 1.5: **Natural gas service areas and the study-area discontinuity**



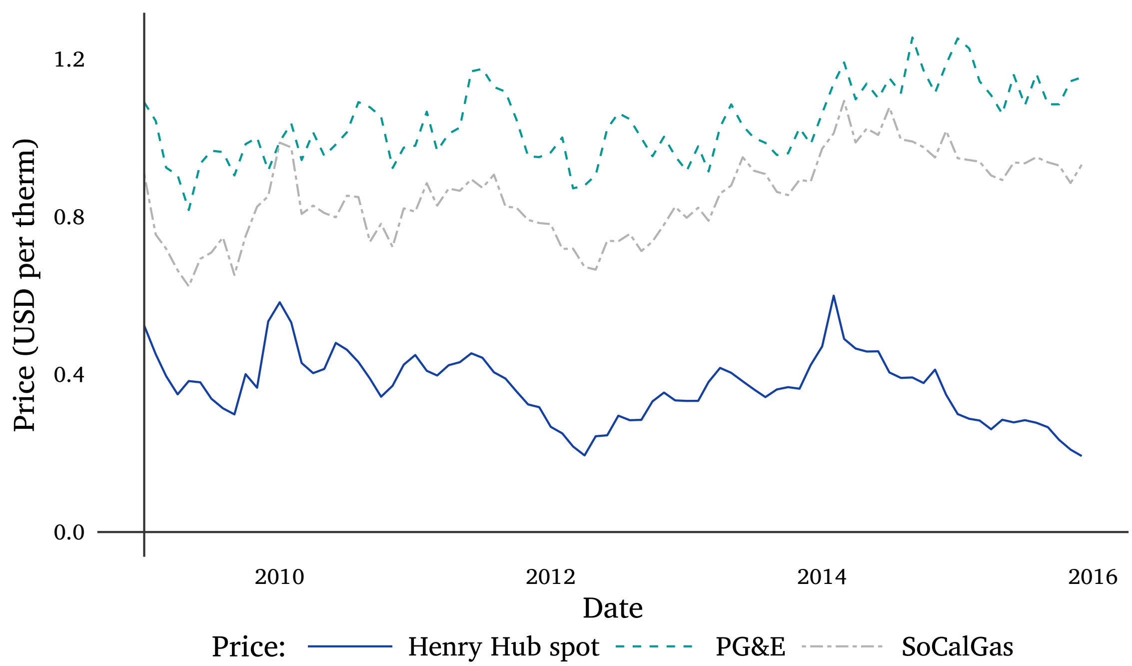
Notes: The left side of the figure displays PG&E's and SoCalGas's services areas (by 5-digit zip code). The right side of the figure zooms in on three clusters of cities that receive service from both utilities. These three clusters of cities encompass 39 zip codes; 18 of these (5-digit) zip codes receive service from both PG&E and SoCalGas. These 18 zip codes represent the main study area for the paper.

Figure 1.6: **Prices across utilities, tiers, and in the spot market**, 2009–2015

(a) **Price regimes over time:** PG&E and SoCalGas, 2009–2015

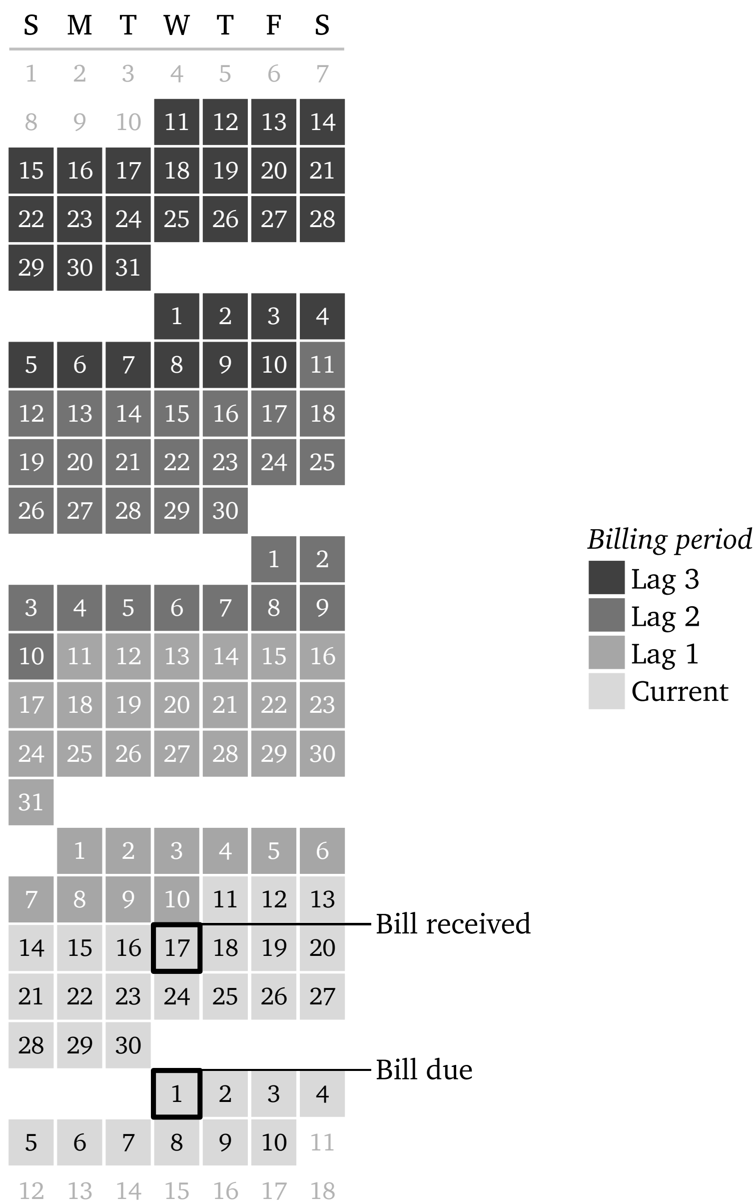


(b) **Correlation across prices** Three relevant natural gas price series, 2009–2015



Notes: *Baseline* refers to first-tier price, *i.e.*, the price a household pays for its first therm of natural gas. *Excess* refers to the second-tier price, *i.e.*, the price a household pays for each therm that exceeds its first-tier allowance (see Figure 1.4). The Henry Hub spot price is generally recognized as a national benchmark (U.S. Energy Information Administration 2016a; Levine, Carpenter, and Thapa 2014).

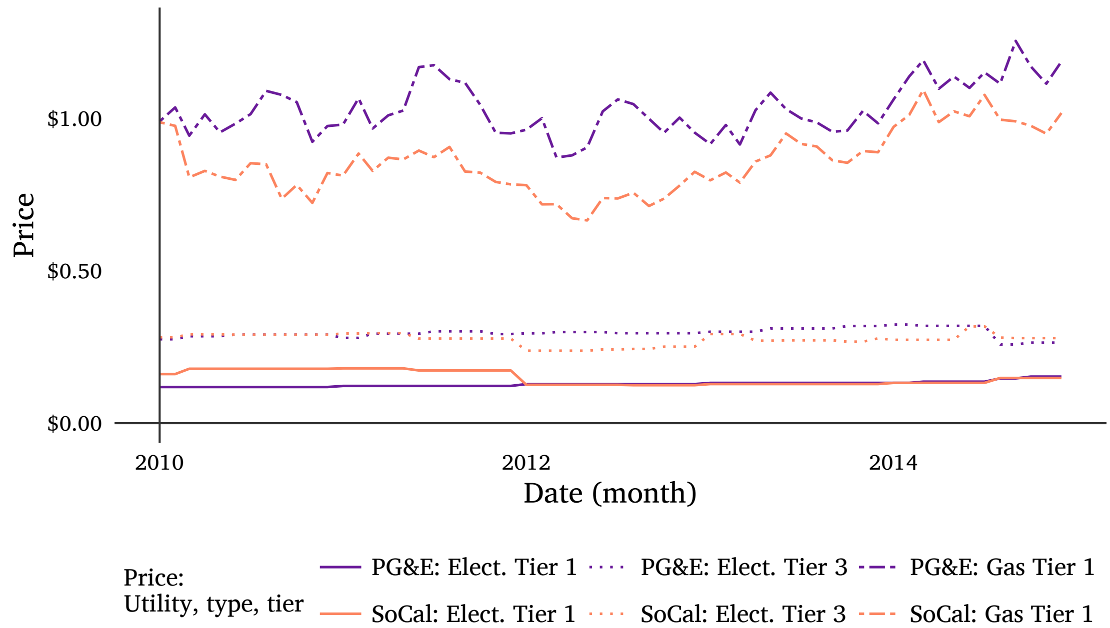
Figure 1.7: **Calendar months and billing periods:** Four 30-day bills and five months



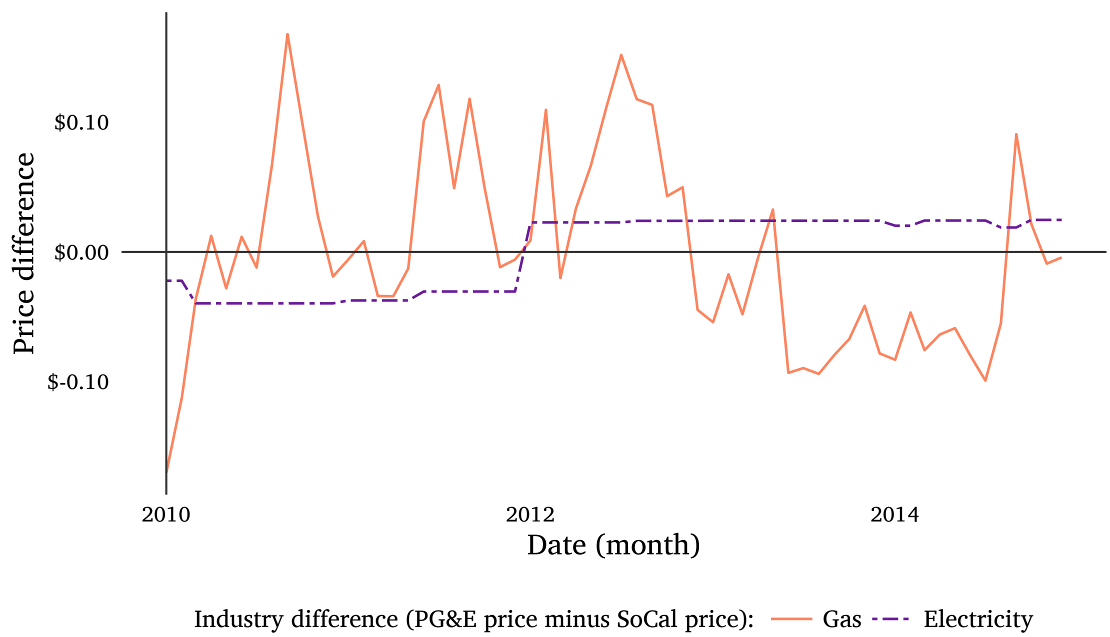
Notes: A consumer receives her bill from the *Lag 1* period on the fifth business day of her *current billing period* (the 17th). Payment for the *Lag 1* bill is due two weeks later (on the 1st). Now consider "Which lag of price is relevant?" **Current:** For the consumer to know the prices in her current billing period, she must read her utility's advice-letter correspondences with CPUC. **Lag 1:** Again, unless she pays attention to her utility's CPUC-approved advice letters, the consumer will not know the prices in the *Lag 1* billing period *until she receives **and** opens her bill*. The bill arrives several days into the new period, and she does not see the bill until payment, the consumer may not learn about the prices of the *Lag 1* bill until the current billing period is nearly complete. Autopay potentially extends this moment of salience even further into the future. **Lag 2:** Throughout the entirety of her *Current* billing period, the consumer knows the prices from her *Lag 2* bill, and for a non-zero amount of time, the *Lag 2* bill is the most recent set of prices the she knows. **Lag 3:** Same level of knowledge as *Lag 2* but less recent.

Figure 1.8: **Natural gas and electricity prices:** Comparing utilities and industries

(a) **Comparing trends in levels**, 2010–2014



(b) **Comparing trends in differences** across utilities, 2010–2014



Notes: For the consumers in this paper, natural gas prices do not significantly correlate with electricity prices—neither in levels (Panel A), nor in differences (Panel B). *Differences* constitute PG&E minus SoCal within the same calendar month. We demean the time series of differences for each industry (natural gas and electricity). *SoCal* denotes the *Southern California Gas Company* for natural gas and *Southern California Edison* for electricity. The underlying data come from publicly available CPUC letters for the relevant utilities.

1.8 Tables

1.8.1 Descriptive tables

Table 1.1: **Prior point estimates:** The price elasticity of demand for residential natural gas

Paper	Data	Estimate
Davis and Muehlegger (2010)	US state panel	−0.278
Maddala *et al.* (1997)	US state panel	−0.09 to −0.18
Garcia–Cerrutti (2000)	Calif. county panel	−0.11
Hausman and Kellogg (2015)	US state panel	−0.11
Herbert and Kreil (1989)	Monthly time series	−0.36
Houthakker and Taylor (1970)	Time series	−0.15
Metcalf and Hassett (1999)	RECS HH panel	−0.08 to −0.71
Meier and Rehdanz (2010)	UK HH panel	−0.34 to −0.56
Rehdanz (2007)	Germany HH panel	−0.44 to −0.63

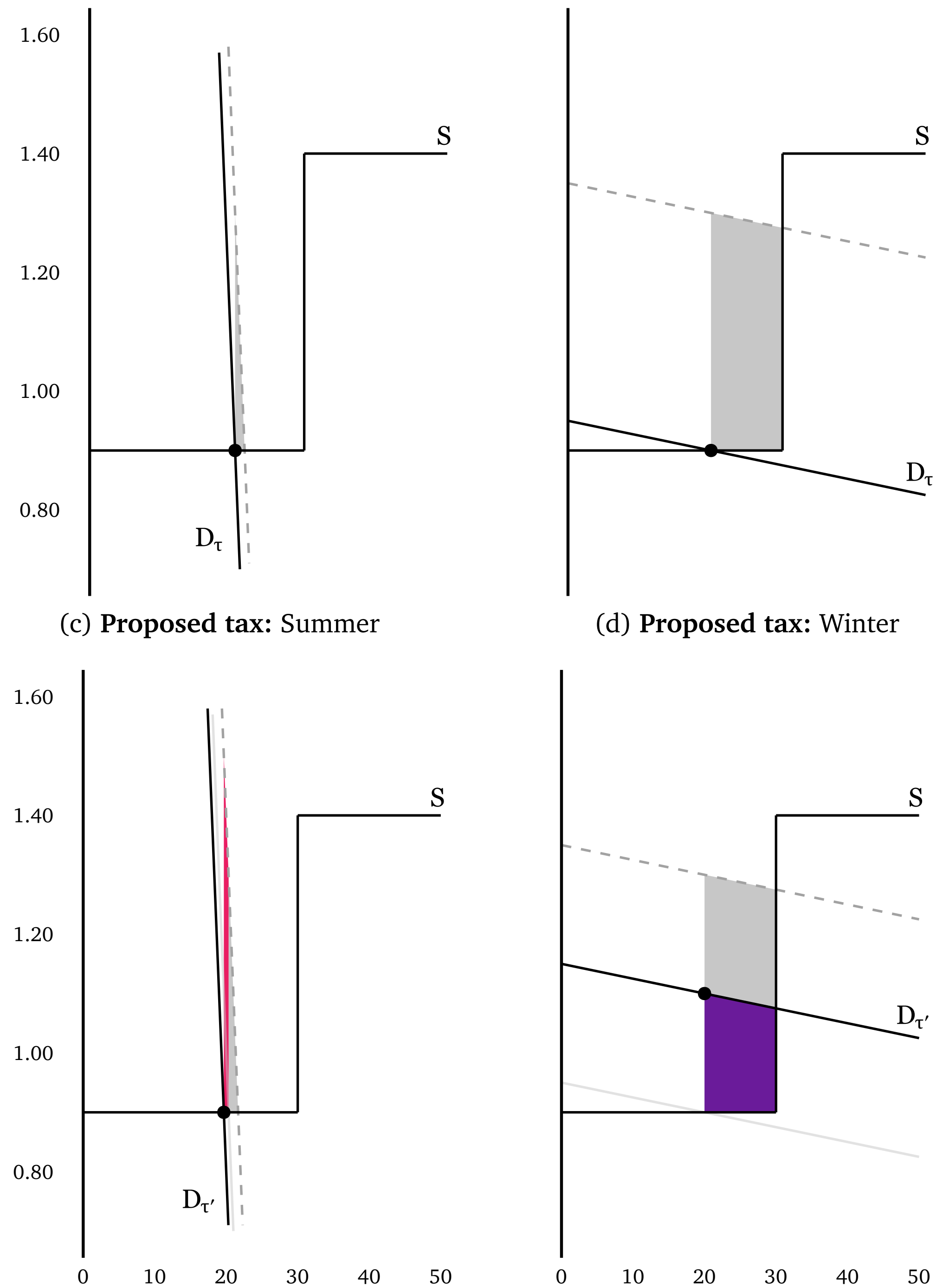
Sources: Authors and Alberini et al. (2011)

Table 1.2: **Billing data summaries**

	Full dataset		Border-area dataset	
	PG&E	SoCalGas	PG&E	SoCalGas
N. 5-digit zip codes	597	611	18	18
N. 9-digit zip codes	680,846	610,207	18,047	16,295
N. unique households	5,888,276	2,526,503	152,418	68,407
N. bills	180,663,705	95,335,393	3,401,947	2,352,141
Approx. value (USD)	\$5.71B	\$3.28B	\$120M	\$70.5M

Notes: *Full dataset* refers to all of the PG&E and SoCalGas bills in the data. *Border-area (discontinuity) dataset* refers to the subset of the *full dataset* for households located in the 18 5-digit zip codes served by both utilities during 2010–2014.

Figure 1.9: **Increasing tax efficiency** using seasonal heterogeneity



(a) **Current tax:** Summer (b) **Current tax:** Winter

(c) **Proposed tax:** Summer (d) **Proposed tax:** Winter

Notes: Each figure presents the combination of a tax (current *vs.* proposed) and a season; the *x* and *y* axes are quantity and price, respectively. The top row illustrates the two seasons under the **current tax**, where households pay the same tax per therm in both seasons. The shaded gray area gives the deadweight loss (DWL) under this tax. **Proposed tax:** The bottom row doubles the tax in the summer—increasing DWL by the narrow pink region—and halves the tax in the winter—reducing DWL by the purple region. Overall DWL decreases.

Table 1.3: **Numerical summaries:** Prices, quantities, and other variables of interest

	5% Sample of California			**Border-discontinuity sample**				
		Split by utility			**Split by season**		**Split by CARE**	
Variable	**Overall**	PG&E	SoCalGas	**Overall**	Winter	Summer	CARE	Non-Care
Baseline price	0.890	0.982	0.743	0.903	0.884	0.920	0.808	0.981
	[0.169]	[0.121]	[0.124]	[0.142]	[0.136]	[0.145]	[0.085]	[0.131]
Average price	1.014	1.105	0.868	1.021	1.001	1.040	0.909	1.115
	[0.185]	[0.144]	[0.144]	[0.162]	[0.158]	[0.164]	[0.100]	[0.143]
Marginal price	1.021	1.128	0.850	1.039	1.012	1.064	0.934	1.126
	[0.226]	[0.186]	[0.173]	[0.198]	[0.191]	[0.202]	[0.145]	[0.194]
Therms	35.463	37.754	31.814	33.827	50.954	17.731	33.114	34.420
	[33.780]	[36.011]	[29.579]	[30.770]	[35.249]	[11.580]	[28.763]	[32.331]
Days	30.399	30.428	30.353	30.399	30.588	30.223	30.404	30.396
	[1.428]	[1.267]	[1.651]	[1.304]	[1.384]	[1.197]	[1.276]	[1.326]
Therms per day	1.159	1.236	1.038	1.106	1.659	0.587	1.084	1.125
	[1.092]	[1.170]	[0.943]	[0.994]	[1.135]	[0.384]	[0.930]	[1.043]
Total bill	36.870	42.394	28.075	34.951	52.075	18.857	30.314	38.804
	[39.576]	[44.056]	[29.045]	[33.881]	[39.897]	[14.007]	[27.257]	[38.102]
(Percent) CARE	27.43%	26.35%	29.15%	45.38%	45.00%	45.74%	100%	0%

Notes: Unbracketed values provide the variables' means; bracketed values denote the variables' standard deviations. The *5% sample of California* is based upon 5% of PG&E's and SoCalGas's natural gas bills from 2010–2014, sampling at the 5-digit zip code. The *border-discontinuity sample* represents all bills from PG&E and SoCalGas for the 18 5-digit zip codes served by both utilities from 2010–2014.

Table 1.4: **Balance on observables:** Comparing utilities' customers across the discontinuity

	Non-CARE			CARE		
Variable	PG&E	SoCalGas	**Diff.**	PG&E	SoCalGas	**Diff.**
Panel A: Summer						
Therms	17.61	17.29	0.32	19.35	18.00	1.34
	[10.8]	[11.7]	[11.3]	[11.3]	[11.3]	[11.3]
Days in bill	30.31	29.97	0.34	30.29	29.96	0.33
	[1.16]	[1.36]	[1.28]	[1.16]	[1.36]	[1.22]
Allowance	14.17	17.22	−3.05	14.14	17.11	−2.96
	[0.805]	[8.05]	[6.14]	[0.851]	[8.17]	[4.33]
Total bill	21.58	16.45	5.14	19.03	13.52	5.51
	[14.8]	[12.4]	[13.8]	[12.4]	[9.35]	[11.9]
HDDs	0.16	0.25	−0.08	0.14	0.26	−0.12
(thousands)	[0.309]	[0.407]	[0.367]	[0.267]	[0.418]	[0.315]
N	810,949	961,824	1,772,773	973,063	320,082	1,293,145
Panel B: Winter						
Therms	51.40	54.07	−2.67	49.60	49.94	−0.34
	[33.8]	[35.7]	[34.8]	[31.1]	[33.1]	[31.6]
Days in bill	30.55	30.78	−0.24	30.57	30.83	−0.26
	[1.31]	[1.8]	[1.59]	[1.31]	[1.81]	[1.45]
Allowance	46.70	49.07	−2.37	47.16	49.68	−2.52
	[12.8]	[10.7]	[11.8]	[12.4]	[10.4]	[12]
Total bill	59.79	50.60	9.19	45.35	36.51	8.84
	[41.8]	[36.4]	[39.4]	[30.3]	[26.5]	[29.7]
HDDs	1.69	1.73	−0.04	1.70	1.75	−0.05
(thousands)	[0.467]	[0.437]	[0.452]	[0.439]	[0.422]	[0.435]
N	746,140	800,037	1,546,177	871,795	270,198	1,141,993

Notes: Unbracketed values provide the variables' means; bracketed values denote the variables' standard deviations. The standard deviations below the difference column (*Diff.*) are pooled across utilities. The difference column denotes the difference in means across utilities for the given cross-section of data. For example, the rightmost **Diff.** column in **Panel A** gives the difference between the PG&E mean and the SoCalGas mean for CARE households in summer months, $\bar{X}_{\text{PGE}} - \bar{X}_{\text{SCG}}$. *CARE* households participate in the California Alternative Rates for Energy (CARE) program. CARE targets low-income households and provides a 20 percent discount on natural gas bills. Heating degree days (HDDs) are in thousands. We calculate the number of heating degrees for day t with mean temperature $\bar{T}_t$ (in °F) as $\text{HDD}_t = \mathbb{1}\left\{\bar{T}_t < 65\right\} \times \left(65 - \bar{T}_t\right)$. The HDDs variable above is thus $\text{HDDS} = \sum_t \text{HDD}_t/1000$.

Table 1.5: **OLS Results**: Estimating elasticities, varying the dataset, price, and fixed effects

	Dependent variable: Log(Consumption, daily avg.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Marginal price)	0.4698***	0.4346***	0.4276***			
	(0.0106)	(0.0136)	(0.0134)			
Log(Baseline price)				0.0217	−0.0918***	−0.1009***
				(0.0147)	(0.0201)	(0.0209)
Bill HDDs	T	T	T	T	T	T
Household FE	T	T	T	T	T	T
Month-of-sample FE	T	T	F	T	T	F
City by month-of-sample FE	F	F	T	F	F	T
Sample	5% CA	Border	Border	5% CA	Border	Border
N	12,855,910	5,754,088	5,754,088	12,855,910	5,754,088	5,754,088

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing cycle. *Base* or *baseline* price refers to the price the household pays for its first unit (*therm*) of natural gas. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Significance levels:* *10%, **5%, ***1%.

1.8.2 2SLS results

Table 1.6: **First- and second-stage results**:
Instrumenting consumers' prices with the Henry Hub spot price

Dependent variable: Log(Consumption, daily avg.)

	(1) Marginal	(2) Average	(3) Avg. Mrg.	(4) Baseline	(5) Sim. Mrg.
Panel A: First-stage results					
Spot price	0.3679***	0.3697***	0.3384***	0.4699***	0.3949***
	(0.0774)	(0.0521)	(0.0570)	(0.0434)	(0.0840)
Spot price × SoCalGas	0.7868***	0.7174***	0.9389***	0.8212***	0.8174***
	(0.0299)	(0.0186)	(0.0198)	(0.0176)	(0.0317)
Panel B: Second-stage results					
Log(Price) *(instrumented)*	−0.2098***	−0.2312***	−0.1734***	−0.2030***	−0.1705**
	(0.0706)	(0.076)	(0.0585)	(0.065)	(0.0698)
First-stage F stat.	418.4	899.4	1,311.0	1,333.2	369.9
Bill HDDs	T	T	T	T	T
Household FE	T	T	T	T	T
City mo.-of-sample FE	T	T	T	T	T
N	5,754,085	5,754,085	5,754,085	5,754,085	4,682,526

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the houesholds' bill. *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Avg.* or *average* price is the total bill divided by quantity. *Avg. Mrg.* or *average marginal* price denotes the quantity-weighted average of the household's marginal price. *Sim. Mrg.* or *simulated marginal* price is the household's marginal price (using the relevant pricing regime) as a function of the household's historical consumption patterns (lagged bills 10 through 14). As discussed in the empirical strategy section, the numbers of observations differ due to the lags required to calculate the *simulated instrument* for marginal price.
Significance levels: *10%, **5%, ***1%.

Table 1.7: **Comparing lags, second-stage results**: Marginal and average prices with HH spot price IV

Dependent variable: Log(Consumption, daily avg.)

	Marginal Price				Average Price			
	(1) 1 Lead	(2) No lag	(3) 1 Lag	(4) 2 Lags	(5) 1 Lead	(6) No lag	(7) 1 Lag	(8) 2 Lags
Log(Price)	0.0480	−0.1121	−0.0223	−0.2098***	0.0515	−0.1244	−0.0177	−0.2312***
instrumented	(0.0902)	(0.0762)	(0.0668)	(0.0706)	(0.0972)	(0.0805)	(0.0730)	(0.0760)
First-stage F stat.	326.7	337.9	410.8	418.4	679.1	725.8	884.4	899.4
Bill HDDs	T	T	T	T	T	T	T	T
Household FE	T	T	T	T	T	T	T	T
City-month FE	T	T	T	T	T	T	T	T
N	5,501,467	5,754,088	5,754,088	5,754,085	5,501,467	5,754,088	5,754,088	5,754,085

Notes: The city-month FE is the interaction of city and month-of-sample. With regard to lags: *No lag* refers to the price for the household's contemporaneous bill; *1 Lag* refers to the price on the household's previous bill; *etc. Avg.* or *average* price is the total bill divided by quantity. *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Significance levels:* *10%, **5%, ***1%.

1.8.3 Heterogeneity results

Table 1.8: **Heterogeneity by season or income**:
Second-stage results, instrumenting marginal price with HH spot price

Dependent variable: Log(Consumption, daily avg.)

	Marginal Price			
	Split by **Season**		Split by **CARE (Income)**	
	(1) Summer	(2) Winter	(3) CARE	(4) Non-CARE
Log(Price)	0.0519^{*}	-0.3769^{***}	-0.2443^{***}	-0.1413^{**}
instrumented	(0.0285)	(0.1399)	(0.0794)	(0.0684)
First-stage F stat.	319.6	174.2	393.7	335.8
Bill HDDs	T	T	T	T
Household FE	T	T	T	T
City month-of-sample FE	T	T	T	T
N	3,065,917	2,688,168	2,435,135	3,318,950

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Summer* includes April through September. *Winter* includes October through March. *CARE* households participate in the California Alternative Rates for Energy (CARE) program. CARE targets low-income households and provides a 20 percent discount on natural gas bills. We estimate the heterogeneity results by splitting the sample along the dimension(s) of heterogeneity and then individually estimating the models. *Significance levels:* *10%, **5%, ***1%.

Table 1.9: **Heterogeneity by season and income**:
Second-stage results, instrumenting marginal price with HH spot price

Dependent variable: Log(Consumption, daily avg.)

	Marginal Price			
	(1) Summer CARE	(2) Summer Non-CARE	(3) Winter CARE	(4) Winter Non-CARE
Log(Price) *instrumented*	0.0457 (0.0353)	0.0742** (0.0324)	−0.5226*** (0.1424)	−0.3173** (0.1498)
First-stage F stat.	303.4	237.1	145.6	156.7
Bill HDDs	T	T	T	T
Household FE	T	T	T	T
City month-of-sample FE	T	T	T	T
N	1,293,144	1,772,773	1,141,991	1,546,177

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Summer* includes April through September. *Winter* includes October through March. *CARE* households participate in the California Alternative Rates for Energy (CARE) program. CARE targets low-income households and provides a 20 percent discount on natural gas bills. We estimate the heterogeneity results by splitting the sample along the dimension(s) of heterogeneity and then individually estimating the models. *Significance levels:* *10%, **5%, ***1%.

2 Are our hopes too high? Testing cannabis legalization's potential to reduce criminalization

Chapter abstract Cannabis legalization advocates often argue that cannabis legalization offers the potential to reduce the private and social costs related to criminalization and incarceration—particularly for marginalized populations. While this assertion is theoretically plausible, it boils down to an empirically testable hypothesis that remains untested: does legalizing a previously illegal substance (cannabis) reduce arrests, citations, and general law-enforcement contact? This paper provides the first causal evidence that cannabis legalization need not necessarily reduce criminalization—and under the right circumstances, may in fact increase police incidents/arrests for both cannabis products and non-cannabis drugs. First, I present a theoretical model of police effort and drug consumption that demonstrates the importance of substitution and incentives for this hypothesis. I then empirically show that before legalization, drug-incident trends in Denver, Colorado matched trends in many other US cities. However, following cannabis legalization in Colorado, drug incidents spike sharply in Denver, while trends in comparison cities (unaffected by Colorado's legalization) remain stable. This spike in drug-related police incidents occurs both for cannabis and non-cannabis drugs. Synthetic-control and difference-in-differences empirical designs corroborate the size and significance of this empirical observation, estimating that Colorado's legalization of recreational cannabis nearly doubled police-involved drug incidents in Denver.

2.1 Introduction

Cannabis legalization advocates often argue that cannabis legalization offers the potential to reduce the costs of criminalization and incarceration for society and individuals—and particularly for marginalized populations who especially suffer from the criminalization of drug offenses. While this assertion is theoretically plausible, it boils down to an empirically testable hypothesis that remains untested in the economic and criminal-justice literatures: does legalizing a previously illegal substance (*i.e.*, cannabis) reduce arrests, citations, and general contact with law-enforcement personnel? This paper presents the first causal evidence that cannabis legalization need not necessarily reduce criminalization—defined as crime-related interactions with law-enforcement personnel—and under the right circumstances, may in fact increase incidents, arrests, and citations for both cannabis products and non-cannabis drugs. To formalize the underlying/competing mechanisms, I create a theoretical model for police enforcement effort and consumer's drug choices. This model highlights the importance of effort constraints, substitution, and incentives when asking whether cannabis legalization will reduce criminalization. Then, to empirically test the hypothesis that cannabis legalization reduces criminal-justice pressures, I create a city-level panel of police-based drug incidents—including the type(s) of drug(s) involved in the incidents. I then show that before legalization, trends in drug-related police incidents for Denver, Colorado matched trends for many other US cities. However, following cannabis legalization in Colorado, drug incidents spike sharply in Denver, while trends in comparison cities (unaffected by Colorado's legalization) remain stable. This spike in drug-related police incidents occurs both for cannabis and non-cannabis drugs. Having established this empirical fact, I then use the temporal variation of Colorado's legalization, in conjunction with two different empirical designs—generalized synthetic controls and difference in differences—to formally test whether Colorado's legalization of recreational cannabis reduced (or increased) cannabis and non-cannabis drug incidents in Denver. The results strongly suggest that legalizing recreational cannabis nearly doubled police-involving drug incidents in Denver. These results provide an important case study in drug-policy, policing, and general public policy with many potential lessons. These lessons are particularly relevant today, as many countries, states, and cities are currently revisiting their policies that govern both cannabis and policing.

2.2 Institutions

Before diving into the theory and empirics of cannabis legalization, I will briefly outline the major events and institutions relevant to this paper's case study: Colorado's legalization and implementation of recreational cannabis. Figure 2.1 depicts several important legal events relevant to marijuana legalization in Colorado. Colorado legalized medicinal marijuana use in November 2000 via Amendment 20 and legalized recreational marijuana use in November 2012 through Amendment 64. Colorado governor John Hickenlooper signed Amendment 64 into law on December 10, 2012 and promptly formed a task force to "consider and resolve a number of policy, legal, and procedural issues" (Colorado 2012a). This task force issued its report to the governor in March 2013, making recommendations on regulatory structure, regulatory financing, taxation, licensing requirements, transition to legalized recreational marijuana, consumers, and criminal law. The task force states that its report "contains a plethora of suggestions for safely growing and processing marijuana, as well as packaging and labeling it. The Task Force proposals also are designed to limit the distribution and consumption of marijuana to persons over

21 years of age within the State of Colorado" (Finlaw and Brohl 2013). On May 28, 2013, Governor Hickenlooper signed numerous bills into law—many related to his task force's recommendations—including bills that changed sentencing for drug crimes, defined a "drug-endangered child" with respect to negligence and abuse, updated penalties for driving under the influence of alcohol or drugs, and levied taxes on retail marijuana (Colorado 2013). On January 1, 2014, approximately one year after Colorado legalized recreational marijuana through Amendment 64, Colorado's first retail marijuana stores opened.

While Amendment 64 heralded a new degree of cannabis legality in Colorado, the amendment and later legislative action placed considerable (legal and criminal) restrictions on marijuana purchases, sales, possession, consumption, and cultivation. Amendment 64 prohibited consumption "that is conducted openly and publicly" and restricted possession and transactions to individuals of at least 21 years-of-age (Colorado 2012b). Additionally, Colorado law restricts purchases to 1 ounce for Colorado residents[1], and recreational marijuana consumers must purchase marijuana from licensed marijuana retail stores. Colorado law also forbids driving under the influence of marijuana. Consequences for breaking these laws range from petty offenses, to misdemeanors and felonies—severity generally increasing with volume of marijuana—and punishments can include incarceration and/or fines (NORML 2018; Denver 2018).

2.3 Model

To motivate and illustrate some of the complexities of cannabis legalization, I now set up a model.

I model the decisions and interactions in this setting using two separate and disjoint groups—(**1**) consumers and (**2**) police officers. The members of each group take governance institutions as given (*i.e.*, they are *takers* of the legal status of various substances, goods, and behaviors). I model police officers at the individual-officer level, and I model aggregate (or representative) drug consumption. I assume the actors maximize their own utilities (maximizing expected utility when facing uncertainty over outcomes). Using this model, I show how institutional changes lead to some clear implications and to some ambiguous changes that depend upon the incentives that members of each group face.

2.3.1 (Potential) Drug consumers

Let us first consider a simple model of drug consumption. In this model, the representative consumer maximizes her expected utility from the consumption of two different, illicit drugs by choosing the optimal quantity of each drug.[2] From con-

[1]Before June 2016, Colorado limited residents of other states to purchasing 0.25 ounces (7 grams). Since June 2016, non-residents and residents both can purchase 1 ounce (Pot Guide 2018).

[2]To draw insights from the general equilibrium and strategic nature of legalization, I ignore price and income effects in this model.

suming quantities x_1 and x_2 of the two drugs, the consumer derives utility $U(x_1, x_2)$. However, for each unit of consumption of each drug (x_i), the consumer faces some probability of arrest (π_i) and punishment (Γ_i). Thus, the consumer's utility maximization takes the form

$$\{x_1^*, x_2^*\} = \arg\max_{x_1, x_2} U(x_1, x_2) - \pi_1\gamma_1 x_1 - \pi_2\gamma_2 x_2 \tag{2.1}$$

with first-order conditions

$$U_1 - \pi_1\Gamma_1 = 0 \tag{2.2}$$
$$U_2 - \pi_2\Gamma_2 = 0 \tag{2.3}$$

where U_i denotes the marginal utility of consumption for drug i.[3] These first-order conditions illustrate the tradeoff at the heart of the consumer's problem—balancing the benefit of consumption with its expected punishment—and are consistent with the seminal work in Becker 1968.[4]

2.3.2 Police officers

Now let us turn to a model of policing. Police officers (*officers* henceforth) maximize their on-job utilities by choosing the levels of effort they exert while enforcing various types of crime. Officers receive utility from apprehending offenders and implicitly bear opportunity costs from the allocation of their effort.

Specifically, an individual officer allocates effort e_i to enforcing crime type i from her total allocation of effort E. For simplicity, I consider only two types of crime. Based upon her effort (*e.g.*, hours) enforcing crime i (e_i) and the amount of crime i (x_i), the officer apprehends $n_i(e_i)x_i$) people who committed crime i. I assume that n_i is increasing in effort—the officer will apprehend more people will more effort, holding the size of population of offenders constant—and n_i is concave with respect to effort (diminishing returns to effort).

In this model, an officer may receive different amounts of utility for apprehending offenders of different crimes. I am agnostic to the specific origin of this *reward*—it may come from officers' preferences, social norms, promotion structures, and/or legal institutions. However, the basic motivation comes from recognizing that both society and municipal administrators likely agree that catching murderers matters more than catching jaywalkers. To integrate this reward into the model, I scale the number of individuals of crime type i apprehended by an officer by γ_i.[5]

[3] The second-order conditions are: $U_{ii} < 0$ for each i and $U_{11}U_{22} - U_{12}U_{21} > 0$.

[4] It is also consistent with Bentham's reasoning on punishment: "Montesquieu perceived the necessity of a proportion between offenses and punishments. Beccaria insists upon its importance. But they rather recommend than explain it; they do not tell in what the proportion consists. Let us endeavor to supply this defect, and to give the principal rules of this moral arithmetic. First rule. *The evil of the punishment must be made to exceed the advantage of the offense.*" (Bentham 1840, p. 100)

[5] Another way to think about the parameter γ_i: γ_i gives the degree of criminalization for action i.

Putting these pieces together, the officer chooses optimal levels of effort

$$\{e_1^*,\ e_2^*\} = \arg\max_{e_1,e_2} \gamma_1 n_1(e_1)x_1 + \gamma_2 n_2(e_2)x_2 \quad \text{such that} \quad e_1 + e_2 = E \tag{2.4}$$

2.3.3 Strategic behavior

I assume that the police officers understand that their enforcement decisions (effort) affects the levels of drug consumption chosen by the consumers. I also assume that the consumers do not consider the effect of their choices on the officers' behavior.

If there are κ police officers, then the total number of individuals apprehended for offense i is $\kappa n_i(e_i)x_i$. If apprehension is equally random for each purchase/consumer, then the probability of apprehension π_i can be written as

$$\pi_i = \frac{\kappa n_i(e_i)x_i}{x_i} = \kappa n_i(e_i) \tag{2.5}$$

Consumer responses

Because officers internalize the effect of their effort on consumers' decisions, I link the two models by substituting the probabilities (π_i) given in (2.5) into the first-order conditions shown in (2.2) and (2.3). This substitution results in the necessary conditions

$$U_1(x_1,\ x_2) - \kappa n_1(e_1)\Gamma_1 = 0 \tag{2.6}$$
$$U_2(x_1,\ x_2) - \kappa n_2(e_2)\Gamma_2 = 0 \tag{2.7}$$

Where e_1 and e_2 are the officer's levels of effort, which I assume are exogenous from the consumer's perspective. Applying the implicit function theorem to the first-order conditions in (2.6) and (2.7) yields[6]

$$\frac{\partial x_i}{\partial e_i} = \mathscr{A} U_{jj}\kappa\Gamma_i n_i' \qquad \textit{Sign}{:}\ (+)(-)(+)(+)(+) \qquad = (-) \tag{2.8}$$
$$\frac{\partial x_i}{\partial e_j} = -\mathscr{A} U_{ij}\kappa\Gamma_j n_j' \qquad \textit{Sign}{:}\ -(+)(U_{ij})(+)(+)(+) \qquad = -\textit{sign}(U_{ij}) \tag{2.9}$$
$$\frac{\partial x_i}{\partial \Gamma_i} = \mathscr{A} U_{jj}\kappa n_i \qquad \textit{Sign}{:}\ (+)(-)(+)(+) \qquad = (-) \tag{2.10}$$
$$\frac{\partial x_i}{\partial \Gamma_j} = -\mathscr{A} U_{ij}\kappa n_j \qquad \textit{Sign}{:}\ -(+)(U_{ij})(+)(+) \qquad = -\textit{sign}(U_{ij}) \tag{2.11}$$
$$\frac{\partial x_i}{\partial \kappa} = \mathscr{A}\left[U_{jj}n_i\Gamma_i - U_{ij}n_j\Gamma_j\right] \quad \textit{Sign}{:}\ (-)(+)(+) - (+)(U_{ij})(+)(+) = (-) - \textit{sign}(U_{ij}) \tag{2.12}$$

where $\mathscr{A} = U_{11}U_{22} - U_{12}U_{21}$, the determinant of the consumer's Hessian (and $\mathscr{A} > 0$). Unsurprisingly, (2.8) and (2.10) suggest that as policing effort or punishment severity decrease for drug i, consumption of good i will increase. More

[6] See the consumer comparative statics section in the appendix for the full derivation.

interestingly—particularly for this paper's context—are the indeterminate signs for the "cross" effects in (2.9) and (2.11). If the two drugs are compliments—meaning $U_{ij} > 0$—then as policing effort or punishment severity decrease for drug j, consumption of drug i increases ($i \neq j$). The result flips when the two drugs are substitutes ($U_{ij} < 0$) rather than complements.[7]

Officer optimization

Because an officer understands that her effort allocation affects consumers' behaviors, based up (2.4) the officer's constrained maximization problem takes the form

$$\mathscr{L} = \gamma_1\, n_1(e_1)\;\, x_1(e1,\, e_2) + \gamma_2\, n_2(e_2)\;\, x_2(e1,\, e_2) + \lambda\,(E - e_1 - e_2) \tag{2.13}$$

which has the first-order conditions

$$\mathscr{L}_{e_1} = \frac{\partial \mathscr{L}}{\partial e_1} = \gamma_1 \left(n_1'\; x_1 + n_1\, \frac{\partial x_1}{\partial e_1} \right) + \gamma_2\, n_2\, \frac{\partial x_2}{\partial e_1} - \lambda = 0 \tag{2.14}$$

$$\mathscr{L}_{e_2} = \frac{\partial \mathscr{L}}{\partial e_2} = \gamma_1\, n_1\, \frac{\partial x_1}{\partial e_2} + \gamma_2 \left(n_2'\, x_2 + n_2\, \frac{\partial x_2}{\partial e_2} \right) - \lambda = 0 \tag{2.15}$$

$$\mathscr{L}_{\lambda} = \frac{\partial \mathscr{L}}{\partial \lambda} = E - e_1 - e_2 = 0 \tag{2.16}$$

2.3.4 Comparative statics

Applying the implicit function theorem to the officer's maximization[8] yields the comparative statics

$$\begin{aligned}
\frac{\partial e_i}{\partial \gamma_i} &= \mathscr{B}\left\{ n_i' x_i + n_i' n_i \mathscr{A}\, U_{jj} \kappa \Gamma_i + n_i \mathscr{A}\, U_{ij} \kappa \Gamma_j n_j' \right\} && (2.17)\\
&= (+)\{(+) + (-) + sign(U_{ij})\}\\
\frac{\partial e_i}{\partial \gamma_j} &= \mathscr{B}\left\{ -n_j \mathscr{A}\, U_{ji} \kappa \Gamma_i n_i' - n_j' x_j - n_j' n_j \mathscr{A}\, U_{ii} \kappa \Gamma_j \right\} && (2.18)\\
&= (+)\{-sign(U_{ji}) - (+) - (-)\}\\
\frac{\partial e_i}{\partial \Gamma_i} &= \mathscr{B} \times (\mathscr{A}\, \kappa n_i') \times \left(\gamma_i n_i U_{jj} - \gamma_j n_j U_{ji} \right) && (2.19)\\
&= (+) \times (+) \times \left[(-) - sign(U_{ji}) \right]\\
&< 0 \quad \text{if} \quad U_{ji} \geq 0\\
&< 0 \quad \text{if} \quad \gamma_j n_j U_{ji} < \gamma_i n_i U_{jj} \;\text{ and }\; U_{ji} < 0\\
&> 0 \quad \text{if} \quad \gamma_j n_j U_{ji} > \gamma_i n_i U_{jj} \;\text{ and }\; U_{ji} < 0
\end{aligned}$$

[7]The sign in (2.12) is technically $(+)[(-)(+)(+) - (+)(U_{ij})(+)(+)]$, which simplifies to the formulation given in (2.12).

[8]Described in detail in appendix section *officer optimization*.

$$
\begin{aligned}
\frac{\partial e_i}{\partial \Gamma_j} &= \mathscr{B} \times \left(\mathscr{A} \kappa n_j'\right) \times \left(\gamma_i n_i U_{ij} - \gamma_j n_j U_{ii}\right) \qquad (2.20)\\
&= (+) \times (+) \times \left[sign(U_{ij}) - (-)\right]\\
&> 0 \quad \text{if} \quad U_{ij} \geq 0\\
&> 0 \quad \text{if} \quad \left|\gamma_i n_i U_{ij}\right| < \left|\gamma_j n_j U_{jj}\right| \ \text{ and } \ U_{ij} < 0\\
&< 0 \quad \text{if} \quad \left|\gamma_i n_i U_{ij}\right| > \left|\gamma_j n_j U_{jj}\right| \ \text{ and } \ U_{ij} < 0
\end{aligned}
$$

where $\mathscr{B}$ denotes the (positive) determinant of the officer's first-order conditions (2.14–2.16) with respect to the exogenous parameters $\{\gamma_i, \Gamma_i, \kappa\}$.

As (2.17–2.20) highlight, the effect of decriminalization (reducing Γ_i) on policing effort is ambiguous—depending both upon the degree of substitution or complementarity of the two drugs (U_{ij}) and the incentives officers receive for enforcing the two offenses (γ_1 and γ_2). When the two drugs are compliments, the comparative statics simplify, and a marginal reduction in the penalty for drug i leads to an unambiguous increase in policing effort for enforcement of drug i and an unambiguous decrease in policing effort for drug j.

Accordingly, the effect of decriminalization of drug i on the number of individuals caught by the police for offense i is also ambiguous. For instance, under the simpler case of complements ($U_{12} > 0$), changing the criminal penalty for drug 1 (Γ_1) changes the amount of consumption of drug 1 by

$$
\frac{dx_1(e_1, e_2)}{d\Gamma_1} = \frac{dx_1}{de_1}\frac{de_1}{d\Gamma_1} + \frac{dx_1}{de_2}\frac{de_2}{d\Gamma_1} \qquad (2.21)
$$

$$
= (-)(-) + (-)(+) = (?) \qquad (2.22)
$$

while the officer's effort-to-citation function changes as

$$
\frac{dn_1(e_1)}{d\Gamma_1} = n_1'(e_1)\frac{de_1}{d\Gamma_1} < 0 \qquad (2.23)
$$

Even upon assuming the structure of the complementarity of the two drugs, the effect of decriminalization is still ambiguous.

Rather than generating testable hypotheses, the main takeaway from this exercise is more abstract. This fairly simple and general model does not guarantee that a reduction in the criminal status of one drug will necessarily reduce the number of offenses enforced by police officers for the decriminalized drug. There is also no guarantee for the effects on the second drug, whose legal status does not directly change. These ambiguous results highlight several complexities of decriminalization/legalization. First, police officers' incentives (γ_i) matter—decriminalizing drugs without changing officers' incentives may induce an increase in officer-recorded offenses. Second, the relationships between the decriminalized and non-decriminalized drugs are important—substitutes *vs.* complements flips the sign of several comparative statics. Third, officers allot an approximately fixed amount of effort (or work a fixed number of hours)—meaning a reduction in incentives for one offense likely means an increase in effort toward another offense.[9] Finally, changes in the number of officers can have important effects—an important point, as many legalization/decriminalization regimes include increases in the number of police officers. Taken together, these dimensions suggest that scenarios potentially exist in which legalization/decriminalization of one drug may in fact increase offenses for the decriminalized drug and/or other drugs.

2.4 Data

The empirics in this paper rely upon publicly available data extract files from the National Incident-Based Reporting System (NIBRS) hosted by the National Archive of Criminal Justice Data (NACJD) the Inter-university Consortium for Political and Social Research (ICPSR).[10] NACJD publishes annual extracts of the full dataset, providing data reported police incidents at four levels: (**1**) incident, (**2**) victim, (**3**) offender, or (**4**) arrestee. This paper uses the incident-level extracts for the years 2000–2015.[11]

NIBRS is part of a larger program (the Uniform Crime Reporting Program or UCR) administered by the U.S. Federal Bureau of Investigation (FBI). The NIBRS data constitute "an incident-based reporting system for crimes known to the police" (NACJD 2017). Consequently, each crime incident that occurs within a law-enforcement jurisdiction that reports to the UCR results in a well-documented data point in the NIBRS dataset. Of particular relevance for this paper: NIBRS records the date, types of offenses, types of drugs involved, and the offender's attributes for each incident. While the NIBRS data do not constitute a representative sample of crime in the United States (NACJD 2017), they include the city of interest for this paper (Denver, Colorado) and many other potential control cities.

2.5 Empirical strategy

As outlined in the introduction, the goal of this paper is to examine the empirical evidence whether Colorado's and Denver's paths toward recreational cannabis legalization indeed reduced the criminalization of cannabis and non-cannabis drug offenses in Denver.

The city-level analysis of the effect of cannabis legalization boils down to a comparison of a single *treated* city—Denver, Colorado—to a large number of *untreated* cities of similar sizes to Denver. Figure 2.2 illustrates such a comparison, comparing the number of drug offenses recorded each week for Denver (in black) to other cities.[12] Figures 2.3 and 2.4 illustrate the same time series but split the offenses into

[9] Admittedly, this result is partially built into the model, rather than a result from the model.

[10] Available through the ICPSR website.

[11] At the time of writing this paper (July 2018), the 2015 extract was the most recent extract.

[12] Figure 2.2 compares Denver to 37 other cities. Denver ranks eighth in drug offenses per week during the pre-legalization period. I therefore take the 7 cities ranked above Denver and the 30

cannabis-related offenses (Fig. 2.3) and non-cannabis-related offenses (Fig. 2.4). The time series of weekly, city-level drug-related offenses in Figures 2.2–2.4 suggest that all types of drug offenses—both cannabis and non-cannabis offenses—increased shortly after Colorado legalized recreational cannabis—particularly after the medley of cannabis-related state-legislature bills in May 2013.

To formally test whether Colorado's legalization accompanied increases in drug offenses in Denver, I implement a (generalized) synthetic-controls estimator (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010, 2011, 2015). In addition, I provide estimates using a more traditional difference-in-differences (*DiD*[13]) framework (Card 1990; Bertrand, Duflo, and Mullainathan 2003; Imbens and Wooldridge 2009).

2.5.1 Synthetic controls

This setting—comparing the time series of a single treated city (Denver) to the time-series of many untreated cities—provides a quintessential environment for synthetic-control methods (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010, 2011, 2015; Xu 2017). In fact, Abadie, Diamond, and Hainmueller motivate their canonical synthetic-control study as "based on the comparison of outcomes between units representing the case of interest, defined by the occurrence of a specific event or intervention that is the object of the study, and otherwise similar but unaffected units. In this design, comparison units are intended to reproduce the counterfactual of the case of interest in the absence of the event or intervention under scrutiny." In the current paper's case, Denver represents the "case of interest"—potentially affected by Colorado's legalization of recreational cannabis—and a host of other major American cities (in states without legalized recreational cannabis) constitute the "unaffected units."

While the pre-legalization trends in Figures 2.2–2.4 appear parallel—bearing evidence toward the parallel-trends identifying assumption in a difference-in-differences framework—I first implement a generalized synthetic control empirical strategy, as given by Xu, before turning turning to the more traditional difference-in-differences framework. The strength of generalized synthetic control methods in this setting is that by employing (linear) interactive fixed effects, one can better control for potential unobserved, time-varying heterogeneity across cities ((Pesaran 2006; Bai 2009; Woolridge 2010; Kim and Oka 2014; Gobillon and Magnac 2016; Shi and Lee 2017; Xu 2017)). The *linear interactive fixed effects* take the form of "unit-specific intercepts interacted with time-varying coefficients" (latent factors and factor loadings) (Xu 2017). Gobillon and Magnac argue that the setting of regional policy analysis is particularly important to consider "more elaborate procedures," such as generalized synthetic controls, due to the greater potential for cross-sectional

cities ranked below Denver. The analysis is robust to how I choose this set of cities. Figure 2.2 omits other cities in Colorado because they are also affected by Colorado's legalization and thus offer a poor control for Denver.

[13] Also known as double difference (*DD*).

dependence in regional policies—relationships due to both spatial proximity and economic proximity.

The generalized synthetic-control (GSC) framework assumes a functional form of

$$y_{it} = \tau_{it}\,\mathbb{1}\{t \in \text{Legalization}\}_{it} + \lambda_i' f_t + \varepsilon_{it} \tag{2.24}$$

where i denotes an individual city, t denotes a specific time period (week of sample), and $\mathbb{1}\{\texttt{foo}\}$ denotes an indicator variable that takes a value of one when `foo` is true. Consequently, y_{it} gives the number of incidents recorded in city i during week-of-sample t, and τ_{it} gives the heterogeneous treatment of effect of cannabis legalization for city i at period t—in terms of the number of additional incidents recorded as a result of legalization. As in (Xu 2017), the term $f_t = [f_{1t}, f_{2t}, f_{rt}]'$ gives an $(r \times 1)$ vector of unobserved common factors, and $\lambda_i = [\lambda_{i1}, \lambda_{i2}, \lambda_{ir}]'$ denotes the $(r \times 1)$ vector of factor loadings for city i.[14]

The generalized synthetic-control estimator for the treatment effect of treated city i in time t is the difference between the observed outcome, *i.e.*, $y_{it}(1)$, and the synthetic control for city i in period t, *i.e.*, $\hat{y}_{it}(0)$. The GSC estimate results from (**1**) estimating (2.24) on all control units, (**2**) estimating the factor loadings for each the treatment by minimizing mean square error in the pre-treatment period(s), and (**3**) using the estimates obtained in (**1**) and (**2**) to calculate $\hat{y}_{it}(0)$ (Xu 2017). The estimated treatment effect for (treated) city i in period t is thus $\hat{\tau}_{it} = y_{it}(1) - \hat{y}_{it}(0)$.

2.5.2 Difference in differences

Now consider the simpler—and more restrictive—difference-in-differences framework. The corresponding difference-in-differences estimating equation is

$$y_{it} = \tau\,\mathbb{1}\{t \in \text{Pre-Legalization}\}_t \times \mathbb{1}\{i = \text{Denver}\}_i + \gamma_i + \delta_t + \varepsilon_{it} \tag{2.25}$$

where, as above, i denotes an individual city, t denotes a specific time period (week of sample), and $\mathbb{1}\{\texttt{foo}\}$ denotes an indicator variable that takes a value of one when `foo` is true. The outcome variable y_{it} records the number of police-involved incidents in city i during week-of-sample t, and τ now represents the average treatment effect of cannabis legalization—the number of additional offenses recorded due to cannabis legalization.[15]

The key identifying assumption underlying a difference-in-differences design is the common trends assumption: in the absence of treatment, the control units and treatment units would have remained on similar trajectories (Angrist and Pischke 2009). The pre-cannabis-legalization trends in Figures 2.2–2.4 suggest that Denver and the 37 non-legalization cities observed similar trends prior to cannabis legalization in Colorado—lending credibility toward the assumption that Denver and these

[14]Note that this general definition nests the more common individual-and-time fixed effects setting wherein $f_{1t} = 1$ and $\lambda_{i2} = 1$ (and $r = 2$).

[15]Note that the city and week-of-sample fixed effects γ_i and δ_t obviate traditional indicators for "treatment unit" and "post-period," respectively.

cities would have remained on comparable trajectories in the absence of Colorado's legalization of recreational cannabis in November of 2012.

In order to more formally the pre-treatment trends, I estimate variants of the equation

$$y_{it} = \beta_1 t + \beta_2 t \times \mathbb{1}\{i = \text{Denver}\} + \gamma_i + \delta_t + \varepsilon_{it} \tag{2.26}$$

for months prior to Colorado's cannabis legalization—where t denotes a time trend. Thus, β_1 gives the time trend *pooled* across all observations, and β_2 denotes any differential time trend in Denver. Consequently, I test whether pre-cannabis-legalization trends differed significantly between Denver and the 37 comparison cities by testing whether β_2 differs significantly from zero.

Table 2.1 presents the ordinary least squares (OLS) estimates for β_1 and β_2 in equation 2.26. All four specifications suggest that, on average, the number of drug incidents was trending down in Denver and its comparison cities. In column (1), I estimate (2.26) with an intercept and without any fixed effects—providing raw trends that do not match the difference-in-differences empirical strategy. The results in column (1) indicate that prior to legalization, Denver was on a very similar trend to the comparison cities—though β_2 is statistically significant in column (1), the difference is less than 1 percent of the *pooled* trend (β_2).[16] Furthermore, if Denver's pre-legalization trend was in fact decreasing more steeply than the comparison cities, then the difference-in-differences estimator is biased away from finding that Colorado's legalization of cannabis *increased* the number of drug incidents in Denver. Once I add city fixed effects (column 2) and time fixed effects (columns 3 and 4), there is no significant difference in trends between Denver and the comparison cities.[17]

Finally, Figure 2.5 illustrates an event study wherein I estimate a "treatment effect" for each month of the sample—*i.e.*, the difference in the number of arrests between Denver and the comparison cities, conditional on city, week-of-year, and year fixed effects. The month preceding Colorado's legalization recreational cannabis (October 2012) is the reference month. Specifically, Figure 2.5 depicts the point estimates and standard errors for the θ_k in the equation

$$y_{it} = \sum_{\substack{k=-76 \\ k\neq -1}}^{41} \theta_k \mathbb{1}\{t \in \text{month}_k\} \times \mathbb{1}\{i \in \text{Denver}\} + \gamma_i + \delta_t + \varepsilon_{it} \tag{2.27}$$

where i indexes a city, t references week of sample, γ_i denotes city fixed effects, and δ_t abbreviates both the year fixed effect and the week-of-year fixed effect.

Figure 2.5 bears further evidence that prior to legalization, Denver and the comparison cities observed similar trends in the the number of drug-related police incidents—the point estimates are near zero for many of the months preceding legalization. However, after legalization (line A in Figure 2.5), the the point estimates jump steeply away from zero—consistent with the increases in drug-related

[16]This trend is also quite small relative the the pre-treatment mean of the dependent variable.
[17]Keeping in mind that failure to reject does not prove null.

incidents depicted in Figures Figures 2.2–2.4. Thus, Figure 2.5 common trends assumption in the difference-in-differences design—and also suggests Colorado's legalization increased .

The common trends depicted in Figures 2.2–2.4, the null results for a difference in trends in Table 2.1[18], and the common-trend evidence from Figure 2.5's event study jointly suggest that the common trends assumption seems quite plausible in this setting.

2.6 Results

Having outlined the empirical strategies and the evidence supporting their identifying assumptions, I now review examine the results of estimating the effect of legalized recreational cannabis on drug-related police incidents.

2.6.1 Synthetic controls

Table 2.2 contains the (generalized) synthetic-control estimates for the average treatment effect (on the treated)—estimating the effect of recreational cannabis legalization on the number of drug-involved police incidents per week as give in equation 2.24. Across three separate specifications, the estimated effect is quite stable and precise with point estimates ranging from 35 to 40 (incidents per week) and standards errors of approximately 2.3.

Each column in Table 2.2 represents a different (generalized) synthetic-control specification. Column (1) applies a more traditional synthetic control specification, omitting any time-varying factors and including only fixed effects—for city and for week of sample. Column (2) includes fixed effects for both city and week-of-sample and four other common factors[19] Finally, column (3) removes column (2)'s insistence on individual and temporal fixed effects, allowing the model to fully and flexibly select up to five factors (again through cross validation). In column (3)'s specification, the cross-validated optimal number of common factors is two. Thus, while I allow the generalized synthetic control specifications to vary, the resulting estimates of the effect of cannabis legalization on Denver's drug-related crime instance remain constant. Using point estimates from column (1) of Table 2.2, the causal interpretation of this effect is: Colorado's legalization of recreational cannabis, on average, increased number of drug-involved police incidents in Denver by 41 incidents each week.[20]

Figures 2.6–2.8 illustrate the goodness of fit for Denver's synthetic control in the pre-period (left of the vertical light grey line) and also simulate what drug-

[18]And bearing in mind that all of the difference-in-differences specifications in this paper use fixed effects.

[19]I choose the number of common factors through cross validation—minimizing mean square prediction error (MSPE). See Xu 2017 for more detail on cross validation and choosing the optimal number of factors.

[20]The 95% confidence interval is [36.3, 45.1].

related police incidents may have looked like had Colorado not legalized recreational cannabis in 2012 (the empty pink circles to the right of the vertical line). The pre-legalization synthetic control for Denver tends to match reality quite well—in all three of the specifications—with a few exceptions approximately two years prior to legalization. As indicated by the large and significant estimated effect of legalization on drug incidents, the counterfactuals provided by the synthetic controls in Figures 2.6–2.8 suggest that the number of drug-related police incidents would have been much smaller in nearly every week after November 2012 had Colorado not legalized recreational cannabis.

In Figures 2.6–2.8, post-legalization (reality) Denver appears move farther away from synthetic-control Denver as time passes—suggesting the treatment (legalization) effect may be increasing with time. Figure 2.9 more clearly presents this trend by plotting each week's estimated treatment effect on the y axis (the difference between reality and the synthetic control in each week) separately for the three synthetic-control specifications. I then fit a locally weighted regression (*loess*) to the estimated treatment effects, resetting the loess at the time of legalization. The depicted point estimates and the slope of the loess lines are all consistent with the effect of legalization on the number of drug-related incidents increasing with time—though the slope appears to have leveled off recently.[21]

In summary, three separate specifications of generalized synthetic-control methods estimate that Colorado's legalization of cannabis increased the number of drug-related police incidents by approximately 40 incidents each week—on a base of approximately 40 incidents. The GSC estimator appears to match Denver's actual path in the pre-legalization period before sharply diverging in the weeks following Colorado's legalization of recreational cannabis.

2.6.2 Difference in differences

Table 2.3 presents the difference-in-differences results from estimating equation 2.25. Each column estimates the same equation with a different outcome variable (measured at the city-week level): column (1) uses the total number of drug-related police incidents; column (2) uses the natural log of the total number of drug-related incidents; column (3) restricts the count to cannabis-related incidents; and column (4) considers only non-cannabis related incidents.

Each of the four columns in Table 2.3 supports the synthetic-control results that Colorado's recreational-cannabis legalization significantly and substantially increased drug-related police incidents in Denver.[22] Column (1) estimates that recreational-cannabis legalization increased the number of drug-related police events by 40.57 [27.39, 53.77] incidents each week. The point estimate in column (2) is approximately 0.90, which implies an increase of approximately 146 percent [79%, 213%][23].

[21] This leveling off may be an artifact of the last few weeks of the dataset.

[22] Table B.1 applies week-of-sample fixed effects—rather than week-of-year and year fixed effects—and the point estimates (and standard errors) are essentially unchanged.

[23] Because the treatment variable is an indicator, the percentage-based interpretation of the log-

Columns (3) and (4) of Table 2.3 decompose the number of drug-related incidents by incidents involving cannabis (column 3) and drug incidents that did not involve cannabis (column 4). The results indicate that legalizing recreational cannabis significantly and substantially increased both the number of cannabis-related police incidents (by 12.01 [4.89, 19.13] incidents each week, on a base of 20.26) and the number of non-cannabis drug-related offenses (by 15.31 [10.57, 20.06] incidents each week, on a base of 16.03). While the magnitudes of both effects are quite large, the magnitude of the increase in non-cannabis drug-related incidents is particularly stark, as the estimated 15-incident increase is on a base of 16 incidents.

Finally, the estimated average treatment effect (from the differences-in-differences design) in column (1) of Table 2.3 (40.57) is very similar to the synthetic-control based estimate in column (1) of Table 2.2 (40.68). This results makes sense, as the synthetic control specification in column (1) of Table 2.2 uses very similar fixed effects to the difference-in-differences estimator in column (1) of Table 2.3—*i.e.*, excluding any common factors and covariates. That said, the other synthetic-control estimates in Table 2.2 do not differ substantially from the difference-in-differences results.

2.7 Discussion and conclusions

First, it is important to frame this paper's results. This paper exclusively considers the effects of recreational cannabis legalization on one type of outcome (the number of drug-related police incidents) in one city (Denver, Colorado). This paper finds significant evidence that Colorado's legalization of recreational cannabis increased the number of drug-related police incidents. However, because this paper considers exactly one city affected by legalization, these results do not necessarily apply to all localities considering legalization. In different cities, under different legal/policing strategies, or with different histories/institutions, recreational cannabis legalization may have different effects on criminalization and drug-related incidents. The results do, however, suggest that the common claim that cannabis legalization will reduce drug-related criminalization is not true in *all* cases. In addition, because the goal of this paper is to test the veracity of this specific criminal/social-justice hypothesis, I do not consider other outcome variables. Consequently, the results in this paper should not be taken as representative of the aggregate or distributional changes in individual or social welfare following Colorado's legalization of recreational cannabis. Finally, it is also worth considering the fact that the observed trends may be due, in part, to a transition between very different equilibria: Colorado was among the first states to legalize recreational marijuana, and moving from a regime of (official) prohibition to legalization likely entails periods of disequilibrium. Thus, one should resist extrapolating too far beyond the case of recreational-cannabis legalization in Colorado and its effects on drug-related police incidents in Denver.

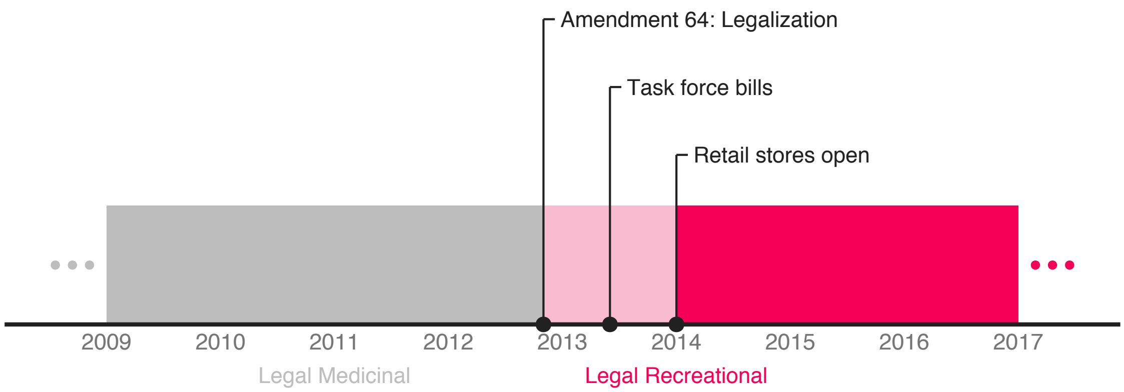
linear specification stems from $\exp\{\hat{\tau}\} - 1$, where $\hat{\tau}$ is the coefficient in column (2) of Table 2.3. I calculate the standard error using the Delta Method.

Caveats made, the empirical evidence in this paper suggests that Colorado's legalization of recreational cannabis causally increased the number of drug-related police incidents. Put simply, rather than decreasing criminalization, cannabis legalization *in this setting* increased drug-related police incidents. Further, the increase in drug-related offenses came from *both* increases in cannabis-involving incidents—which were governed by new laws and citations that developed following Coloradans passed Amendment 64—and increases in non-cannabis drug incidents. While these results may at first blush seem surprising, they are consistent with a model in which police officers respond to (**1**) citizens' drug-consumption choices and (**2**) incentives set by society/local government. The model laid out at the beginning of this paper suggested legalization may increase or decrease incidents involving the legalized and non-legalized drugs. The empirical results in this paper confirm the possibility that legalization may increase criminalization/police incidents for legalized and non-legalized drugs. Future work will likely uncover cases in which legalization has differing effects. However, regardless of future outcomes, this paper demonstrates an important lesson for policymakers:[24] legalization has the potential to increase police-involving incidents—both for the legalized drug and for other drugs. This lesson has important implications for evaluating the effects and equity of policies ranging legalization to criminal prosecution to policing—and is particularly salient today, when inequality, policing, and drug policy weigh heavy in the public conscience.

[24]Including elected officials and voters.

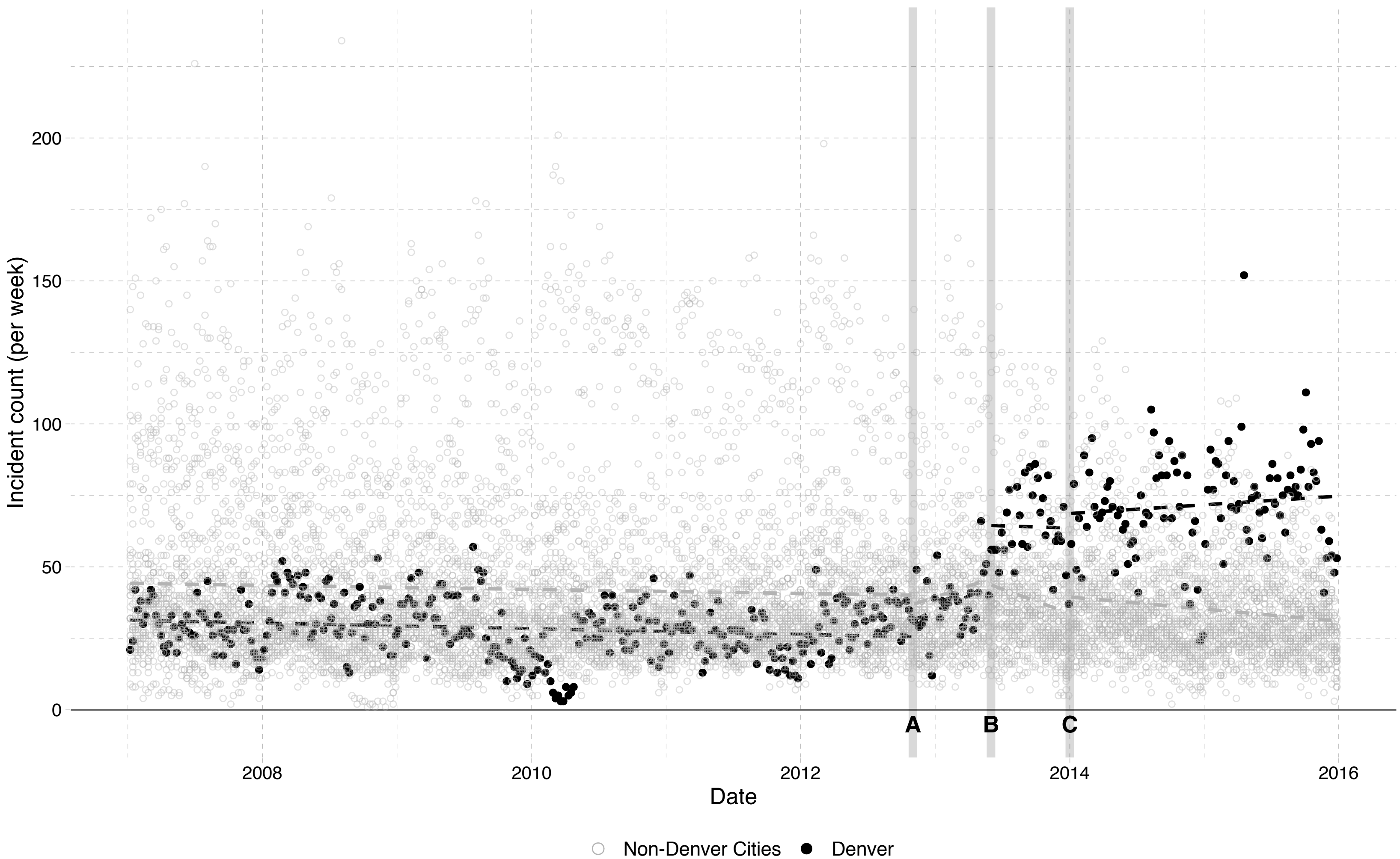
2.8 Figures

Figure 2.1: **Colorado legalization timeline:** Important events for Colorado's implementation of recreational cannabis legalization



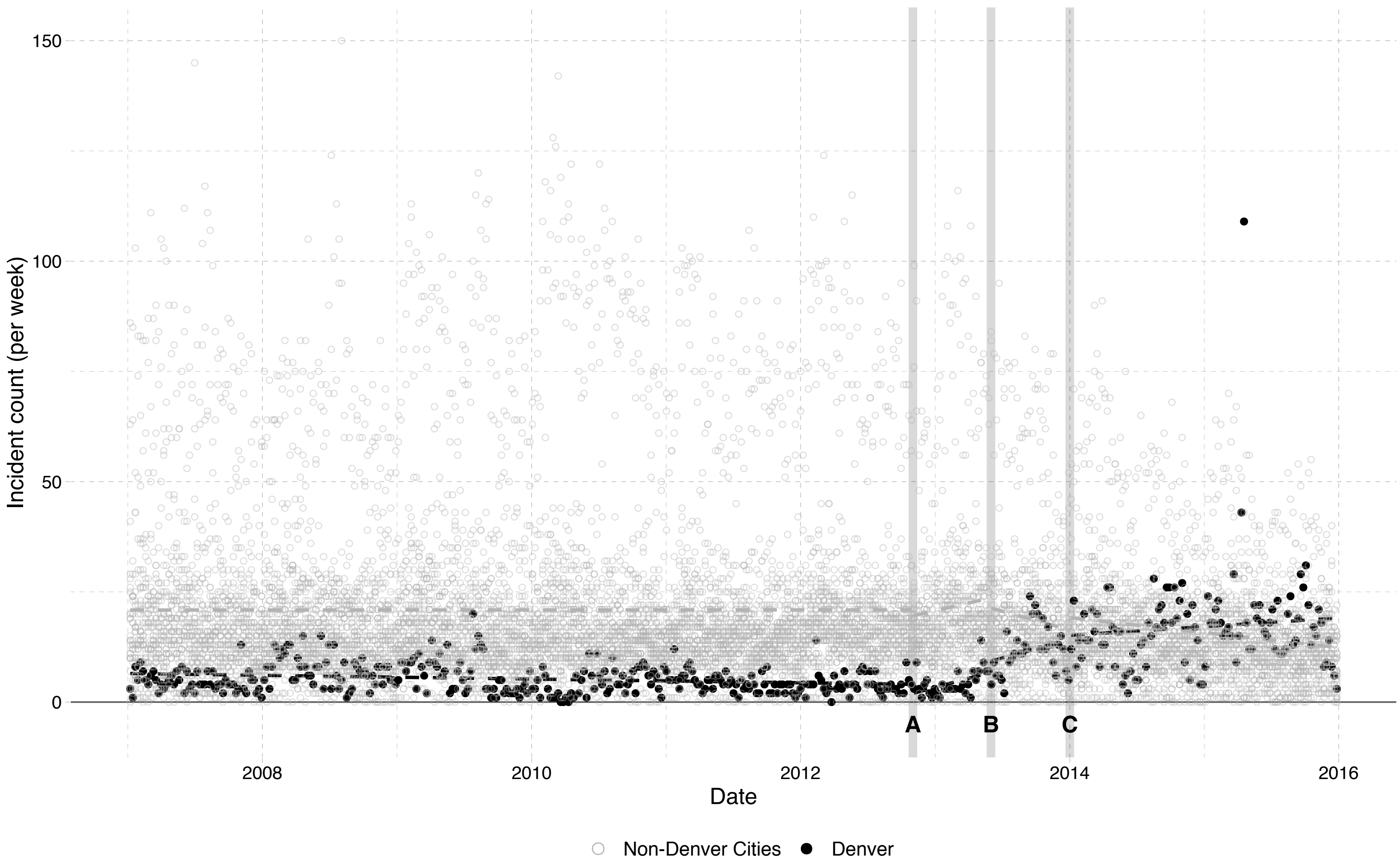
Notes: Coloradans passed Amendment 64 on November 6, 2012, authorizing the legalization of recreation cannabis (with 55% support). On May 28, 2013, Colorado Governor John Hickenlooper signed into law several state-legislature bills that regulate cannabis in Colorado. Recreational cannabis stores open on January 1, 2014.

Figure 2.2: **All drug-related incidents (NIBRS):** Weekly by city, 2007–2016



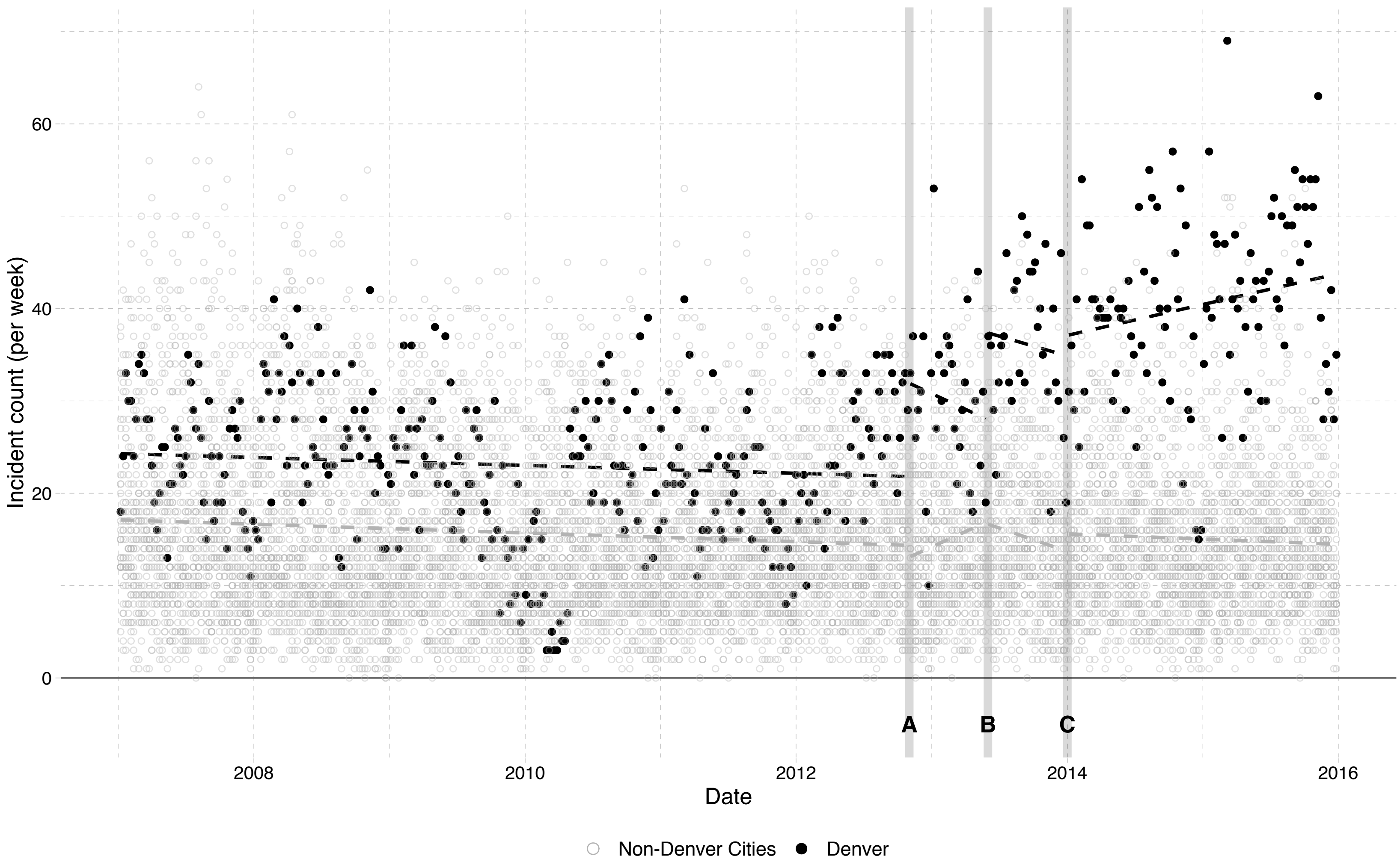
Notes: The lines match the events described in Figure 2.1. Specifically: **A** denotes Colorado's legalization of recreation cannabis via Amendment 64 (November 6, 2012). **B** denotes Governor Hickenlooper signing into law several state-legislature bills that regulate cannabis in Colorado (May 28, 2013). **C** denotes the opening of recreational cannabis stores in January 1, 2014. *Source:* NACJD

Figure 2.3: **Cannabis-related incidents (NIBRS):** Weekly by city, 2007–2016



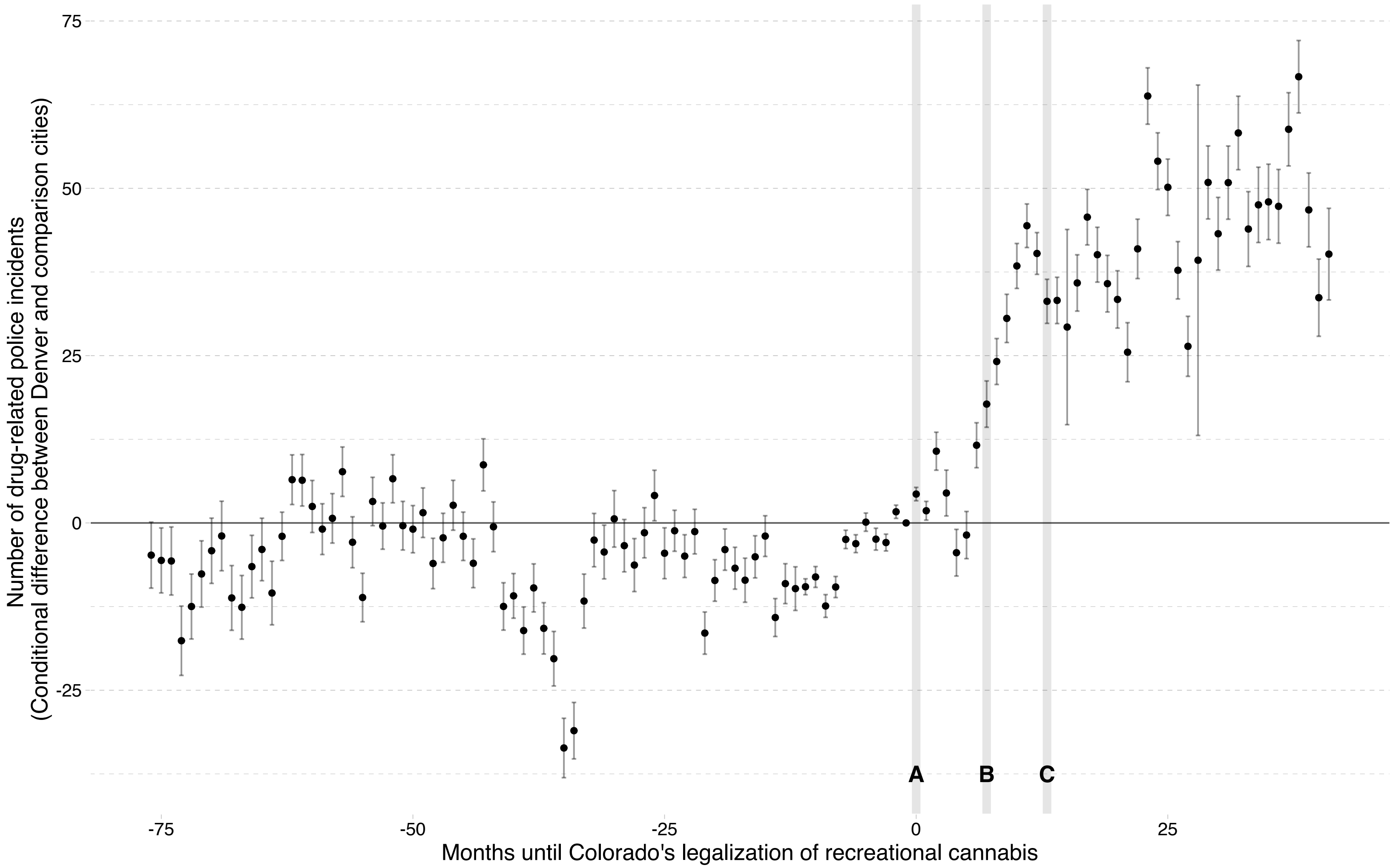
Notes: The lines match the events described in Figure 2.1. Specifically: **A** denotes Colorado's legalization of recreation cannabis via Amendment 64 (November 6, 2012). **B** denotes Governor Hickenlooper signing into law several state-legislature bills that regulate cannabis in Colorado (May 28, 2013). **C** denotes the opening of recreational cannabis stores in January 1, 2014. *Source:* NACJD

Figure 2.4: **Non-cannabis drug-related incidents (NIBRS):** Weekly by city, 2007–2016



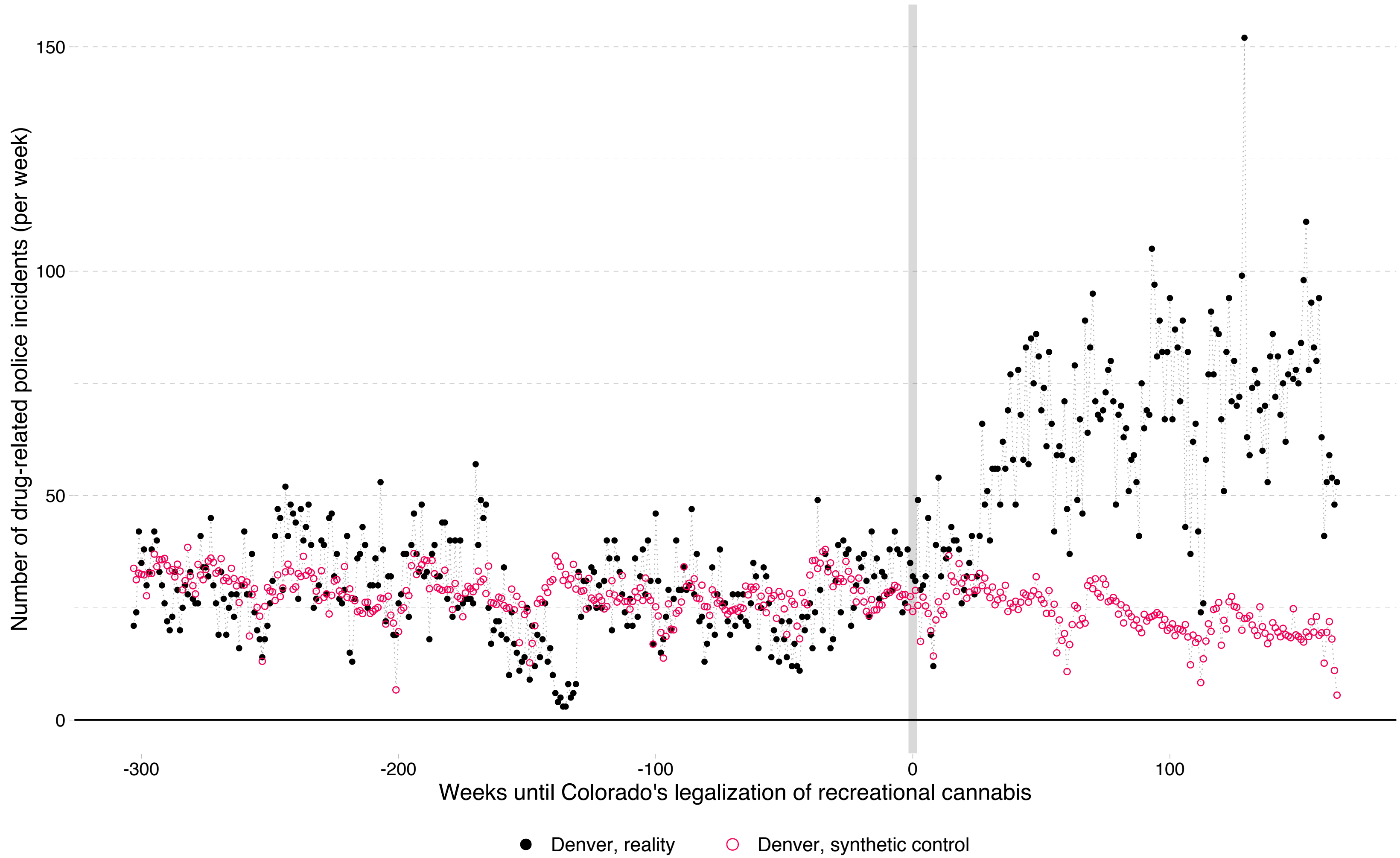
Notes: The lines match the events described in Figure 2.1. Specifically: **A** denotes Colorado's legalization of recreation cannabis via Amendment 64 (November 6, 2012). **B** denotes Governor Hickenlooper signing into law several state-legislature bills that regulate cannabis in Colorado (May 28, 2013). **C** denotes the opening of recreational cannabis stores in January 1, 2014. *Source:* NACJD

Figure 2.5: **Event study**: Difference in drug-involving incidents between Denver and comparison cities



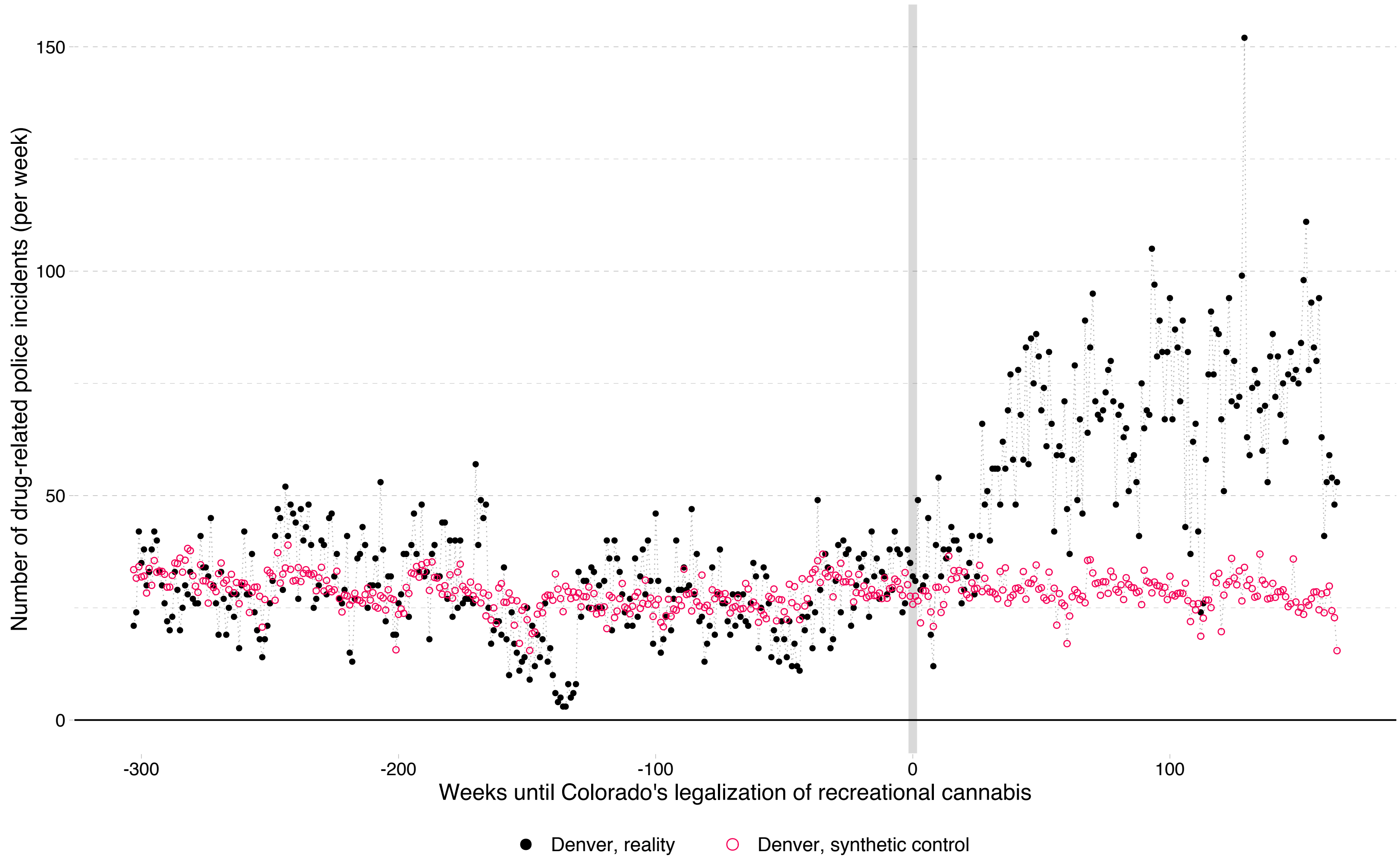
Notes: The lines match the events described in Figure 2.1. Specifically: **A** denotes Colorado's legalization of recreation cannabis via Amendment 64 (November 6, 2012). **B** denotes Governor Hickenlooper signing into law several state-legislature bills that regulate cannabis in Colorado (May 28, 2013). **C** denotes the opening of recreational cannabis stores in January 1, 2014. The event-study regression includes city, week-of-year, and year fixed effects. I cluster the errors at the city-by-year level. *Source:* NACJD

Figure 2.6: **Synthetic controls, specification 1**: Comparing Denver in reality to synthetic-control Denver



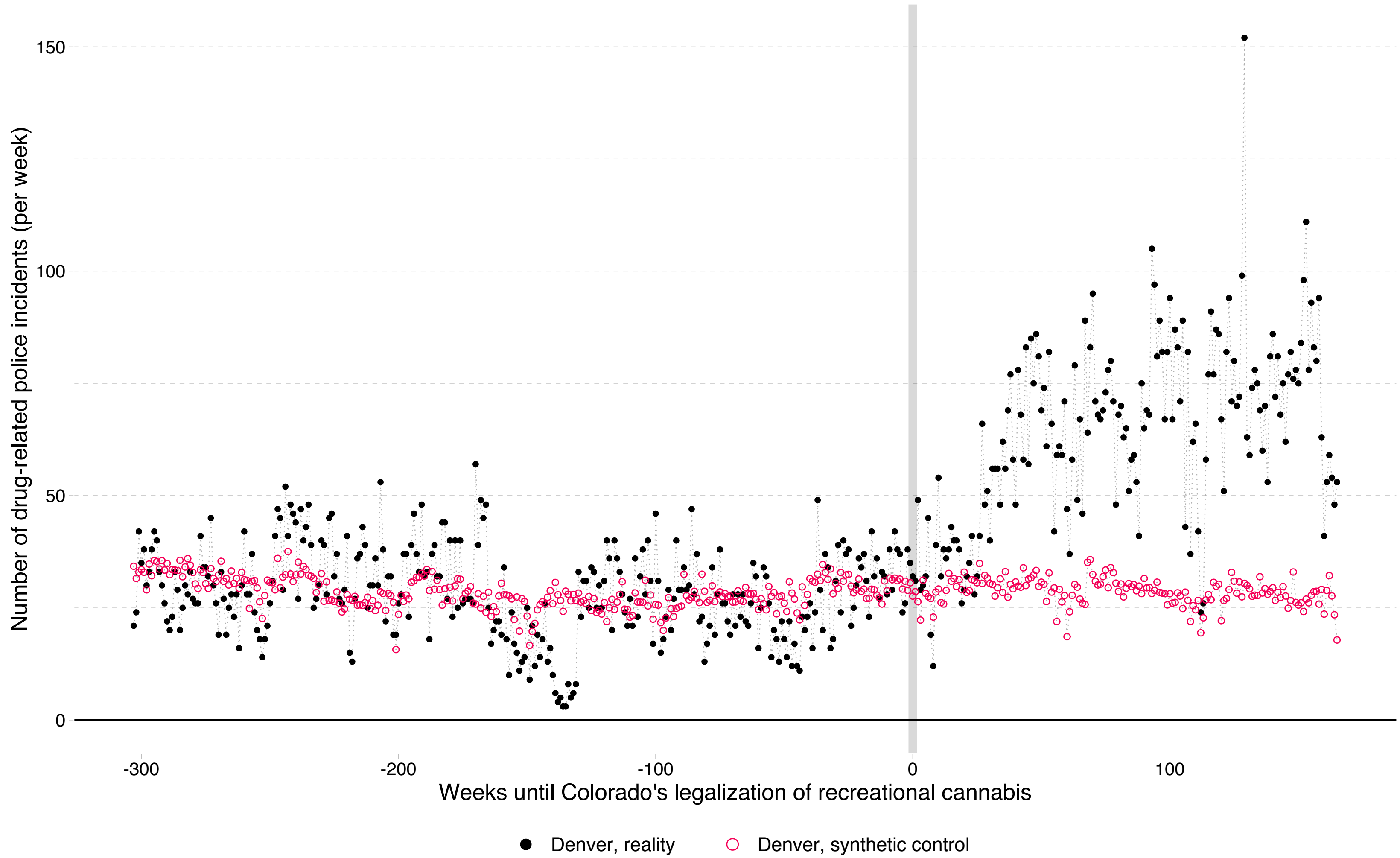
Notes: This figure compares the observed weekly drug-related incidents in Denver (black dots) with the synthetic control for Denver (pink, empty circles), composed of 24 comparison cities. *Specification 1* refers to the synthetic-control specification in column (1) of Table 2.2, which uses city and week-of-sample fixed effects, omitting any further common factors.

Figure 2.7: **Synthetic controls, specification 2**: Comparing Denver in reality to synthetic-control Denver



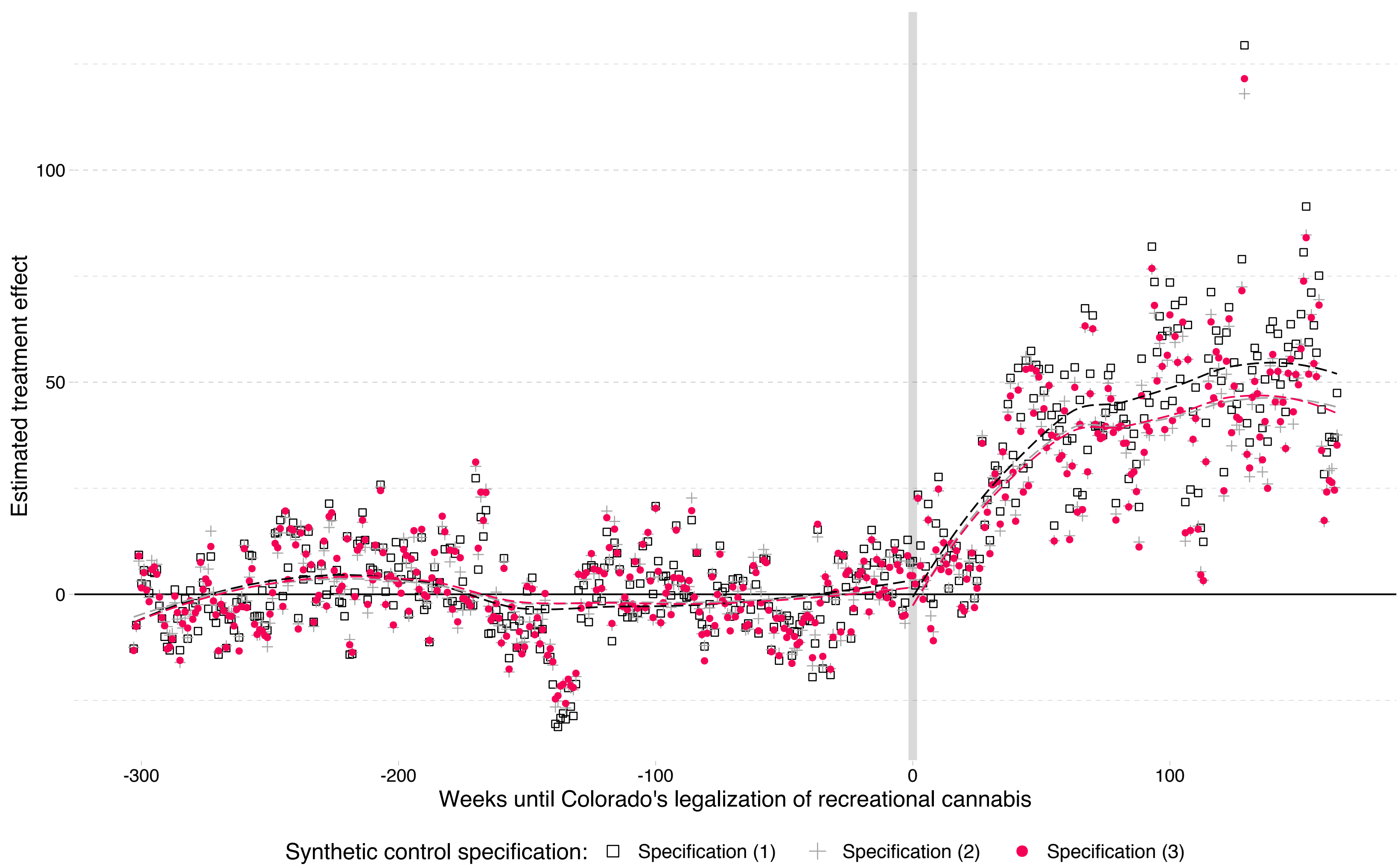
Notes: This figure compares the observed weekly drug-related incidents in Denver (black dots) with the synthetic control for Denver (pink, empty circles), composed of 24 comparison cities. *Specification 2* refers to the synthetic-control specification in column (2) of Table 2.2, which uses city and week-of-sample fixed effects *and* four common factors.

Figure 2.8: **Synthetic controls, specification 3**: Comparing Denver in reality to synthetic-control Denver



Notes: This figure compares the observed weekly drug-related incidents in Denver (black dots) with the synthetic control for Denver (pink, empty circles), composed of 24 comparison cities. *Specification 3* refers to the synthetic-control specification in column (3) of Table 2.2, which uses two common factors (not necessarily using fixed effects).

Figure 2.9: **Synthetic-control treatment effects**: Comparing estimated treatment effects across the three specifications



Notes: This figure compares the estimated treatment effect in each week from the three synthetic-control specifications (using the specification numbers defined in Table 2.2). The dashes depict the loess-smoothed lines—providing local averages of the treatment effects.

2.9 Tables

Table 2.1: **Parallel pre-trends**: Testing pre-legalization, linear trends with OLS

Dependent variable: Weekly drug incidents, (NIBRS)

	(1)	(2)	(3)	(4)
Time trend	-0.6803	-0.6640^{*}		
	(1.26305)	(0.354)		
Time trend	-0.0068^{***}	-0.2849	-0.3154	-0.3166
$\times$ Denver	(0.00135)	(0.937)	(0.847)	(0.869)
Intercept	T	F	F	F
City FE	F	T	T	T
Year FE	F	T	F	F
Week FE	F	F	T	F
Year-Week FE	F	F	F	T
N	7,598	7,598	7,598	7,598

Notes: Each column denotes a separate regression. The observational unit is city i in month of sample t. Errors are clustered within a city in a year. The mean of the dependent variable (the number police-involving drug incidents within a city in a week) is 41.62 in the pre period. *Significance levels:* *10%, **5%, ***1%.

Table 2.2: **Effect of cannabis legalization on drug-related offenses**: Synthetic controls results

Dependent variable: Weekly drug incidents (NIBRS)

	(1)	(2)	(3)
Legalization	40.6833***	35.4370***	35.3196***
ATT	(2.26)	(2.62)	(2.01)
Number of factors	0	4	2
City FE	T	T	F
Week-of-sample FE	T	T	F
N	11,723	11,723	11,723

Notes: The point estimates denote the average treatment effect on the treated. Standard errors result from parametric bootstrapping with n = 1,000. Each column denotes a separate synthetic-control estimation. The specification in (1) includes city and week-of-sample fixed effects—and excludes any other common factors. The results in (2) include both sets of fixed effects and four other common factors. The specification in (3) does not enforce the fixed effects (which are nested in common factors); (3) uses two common factors. I determined the 'optimal' number of common factors in (2) and (3) through cross validation (see Xu 2017). *Significance levels:* *10%, **5%, ***1%.

Table 2.3: **Effect of cannabis legalization on drug-related offenses**: Difference-in-differences results

	Dependent variable			
	(1) N. incidents	(2) Log(N. incidents)	(3) N. cannabis incidents	(4) N. non-cannabis incidents
Legalization *indicator*	40.5779*** (6.73)	0.8994*** (0.14)	12.0104*** (3.63)	15.3135*** (2.42)
City FE	T	T	T	T
Year FE	T	T	T	T
Week FE	T	T	T	T
N	11,723	11,723	11,723	11,723

Notes: Each column denotes a separate regression with a different dependent variable. The observational unit is city i in week of sample t. Errors are clustered within a city in a year. The mean number of drug-related police incidents before Colorado's legalization was 41.62 (20.26 cannabis-related offenses and 16.03 non-cannabis related offenses). *Significance levels:* *10%, **5%, ***1%.

3 | Do aerially applied pesticides affect local air quality? Empirical evidence from California's San Joaquin Valley

Chapter abstract: Many policymakers, public-health advocates, and citizen groups question whether current pesticide regulations properly equate the marginal social costs of pesticide applications to their marginal social benefits—with particular concern for negative health effects stemming from pesticide exposure. Additionally, recent research and policies in public health, epidemiology, and economics emphasize how fine particulate matter (PM2.5) concentrations harm humans through increased mortality, morbidity, mental health issues, and a host of socioeconomic outcomes. This paper presents the first empirical evidence that aerially applied pesticides increase local PM2.5 concentrations. To causally estimate this effect, I combine the universe of aerial pesticide applications in the five southern counties of California's San Joaquin Valley (1.8M reports) with the U.S. EPA's PM2.5 monitoring network—exploiting (**1**) spatiotemporal variation in aerial pesticide applications and (**2**) variation in local wind patterns. I find significant evidence that (upwind) aerial pesticide applications within 1.5km increase local PM2.5 concentrations. The magnitudes of the point estimates suggest that the top decile of aerial applications may sufficiently increase local PM2.5 to warrant concern for human health.

3.1 Introduction

Recent economic, epidemiological, and public health research strengthens the body of evidence that exposure to high levels of pesticides increases the incidence of a number of negative health outcomes, *e.g.*, low birthweight, gestational length, birth abnormalities (Larsen, Gaines, and Deschênes 2017). In addition, a large body of work demonstrates the effect of exposure to particulate matter (PM) on mortality (Seaton et al. 1995; Pope III et al. 2002; Guaita et al. 2011; Lu et al. 2015), morbidity (Pope III 1989; Currie and Walker 2011; Laumbach and Kipen 2012), mental health (Graff Zivin and Neidell 2013; Volk et al. 2013), and negative social/economic outcomes (Ransom and Pope 1992; Kim, Kabir, and Kabir 2015; Isen, Rossin-Slater, and Walker 2017).

This paper attempts to provide the first empirical evidence on (**1**) the extent to which aerially applied pesticides increase local fine particulate matter (PM2.5)[1] concentrations and (**2**) the degree to which aerial pesticide-induced PM2.5 drifts to neighboring areas. To carry out this test, I use the universe of aerial pesticide applications in the five southern counties of California's San Joaquin Valley from 2000 to 2015, in conjunction with the U.S. Environmental Protection Agency (EPA)'s PM2.5 monitoring network. The empirical tests use (**1**) spatiotemporal variation in aerial pesticide application and (**2**) variation in wind direction to isolate the necessary plausibly exogenous variation to evaluate these questions. I find significant evidence that aerially applied pesticides increase local PM2.5 concentrations (reducing local air quality) within 1.5 kilometers of the application. The magnitude of the effect suggests that the top decile of aerially applied pesticides may increase local PM2.5 concentrations sufficiently to warrant public health concern.

While many issues relating to pesticides—health effects, use, bans—remain hotly debated (Pimentel et al. 1992; Tong 2018), federal (*e.g.*, the EPA) and state (*e.g.*, California's Department of Pesticide Regulation (DPR)) entities regulate, to some degree, the types and amounts of pesticides a farmer may apply to her land—particularly in areas near schools or population centers. Notably, the vast majority of pesticide-use regulations consider the individual farmer as the relevant actor—attempting to limit others' exposures from a farmer's pesticide application. This regulatory strategy may be plausibly efficient/optimal if pesticides do not travel far from their points of application and if neighboring farmers' applications do not correlate positively in time. However, if pesticides drift from their points of application—and if farmers tend to apply pesticides at the same time and in the same area[2]—then regulating individual farmers without concern for local, aggregate behavior may miss important dimensions of exposure. In other words, if each farmer in a highly agricultural area applies pesticides just below an established *safe level*, and if each farmer's pesticides aggregate locally and/or drift slightly downwind onto neighboring areas, then these downwind, neighboring areas may be exposed to levels of pesticide above the established *safe level*. This paper investigates the extent to which statistical evidence supports these hypotheses.

3.2 Data

The empirics in this apply three separate datasets: (**1**) fine particulate matter (PM2.5) monitoring data from the EPA's network of air-quality monitors, (**2**) pesticide-use reports from the California Department of Pesticide Regulation (DPR), and (**3**) wind data from NASA's North American Land Data Assimilation Systems, version 2 (NLDAS-2).

[1]Fine particulate matter, or PM2.5, is defined as any particle with an aerodynamic diameter less than 2.5 μm (Volk et al. 2013).

[2]As one might expect if agriculture is correlated in space.

3.2.1 Air-quality monitors

This paper sets out to measure the effect of aerial pesticide applications on local air quality in California's San Joaquin Valley. The ideal candidate for air-quality measurement would measure air quality throughout the southern San Joaquin Valley with high precision and accuracy—and with high temporal frequency. While many valuable measures of air quality exist, as discussed above, a large literature demonstrates the importance of PM2.5 levels for human health. The EPA's PM2.5 monitoring network therefore provides a sensible candidate for this task: between 2000 and 2015, the EPA monitored PM2.5 levels using 53 unique monitors located throughout the five counties of the southern San Joaquin Valley.[3] Further, in nearly every year, at least one monitor measured PM2.5 concentrations on each day of the year, resulting in approximately 108,000 observations of PM2.5 levels in the southern San Joaquin Valley between 2000 and 2015. Table 3.1 summarizes the number of monitors and days observed for each year starting in 2000 and ending in 2015 for the five southern San Joaquin Valley counties. In addition, EPA monitors play a central role in implementing the Clean Air Act in the United States and thus supply a policy-relevant candidate for air-quality measurement in this paper (U.S. E.P.A. 2016).

Each of the five counties on the southern end of California's San Joaquin Valley contain multiple EPA monitors.[4] Figure 3.1 maps out the EPA monitors' locations throughout the five counties. These monitors report hourly PM2.5 local concentrations ($\mu g/m^3$) for a 24-hour period—either each day or every sixth day. Figure 3.2 plots the daily mean PM2.5 reading at each of the 53 monitors on each day the monitor recorded values.[5] Figure 3.2 also illustrates (**1**) substantial seasonal variation in PM2.5 concentrations at these monitors—ramping up in the late fall and peaking in January—and (**2**) the tendency for PM2.5 concentrations in these five counties to exceed established air-quality standards (U.S. E.P.A. 2013). Jointly, Figure 3.1 and 3.2 suggest the EPA PM2.5 network offers a reasonable solution for measuring air quality in the southern San Joaquin Valley, given its spatial and temporal coverage/variation and its importance to environmental regulation.

3.2.2 Pesticide-use reports

The geographic focus of this paper—the southern counties of California's San Joaquin Valley—stems in part from availability of data on pesticide use. In 1990, the state of California established the United States' first state-level, mandatory full-reporting system for pesticides (California D.P.R. 2000). Today the California DPR's pesticide-use reporting (PUR) system is widely regarded as the world's most comprehensive and high-quality record of pesticide use (California D.P.R. 2000;

[3] A number of the monitors phased in and out during this time so 22-36 monitors actively observed each year.

[4] By name, the five counties are: Fresno County, Kern County, Kings County, Madera County, and Tulare County.

[5] Figure C.1 repeats this exercise but uses the daily *maximum* PM2.5 instead of the *mean*.

Wilhoit 2012; Larsen, Gaines, and Deschênes 2017). By law, any individual who applies agricultural pesticides must report the pesticide applications on a monthly basis to the relevant county agricultural commissioner. The county agricultural commissioner then sends the reports to the California DPR, who review, summarize, and publish the PUR data.[6] For application of agricultural pesticides, the user must report (**1**) the date of application, (**2**) the location of application (section, township, range), (**3**) the type of pesticide, (**4**) the amount of pesticide, and (**5**) identifiers for the site and user (California D.P.R. 2000).[7] Because the PUR system does not systematically identify PURs below the section-township-range level, the finest spatial resolution is the section, which is a grid of approximately one-square-mile cells. This spatial restriction both limits attribution of pesticide use to PM2.5 concentrations and limits the usefulness of wind-direction data, as I discuss in more detail below.

Figures 3.3 and 3.4 depict the number of pesticide use reports and the tons of pesticides applied, respectively, aggregated to the month of sample—across the five study counties from 2000 to 2015. Each figure splits the summaries by (a) aerially applied pesticides and (b) ground-applied pesticides. Figure 3.5 illustrates the amount (number of tons) of pesticides applied by day of week across the five counties in the same time period.[8] The three sets of figures emphasize suggest several relevant points. First, the five study counties are very active in pesticide application—in the number of applications and in the amount (tons) of pesticides applied. Second, time trends are quite apparent—both in annual cycles and weekly cycles. Third, the time trends differ by the type of pesticide application (aerial versus ground) and by the measure of application (count of PURs versus tonnage of pesticides). Importantly, these time trends—or temporal clustering—exactly describe a situation in which one may be concerned about the build up of particulate matter from pesticide aggregation or drift.

Figures 3.6 and 3.7 map the intensities of aerial and ground pesticide use, respectively—summing the total amount of pesticides applied within each section from 2000 to 2015. The color scale of the sections' shading depicts the log[9] intensity of pesticide use within the section over the 16-year period. The white dots denote school locations[10] to visually proxy for human population, and the white lines delineate the five counties' borders. Both maps illustrate the spatial variability of pesticide use in California—and the intensity in many locations. The two maps in Figures 3.6 and 3.7 also highlight the proximity of high levels pesticide applications to human populations. While there are holes in density of pesticide applications corresponding to the locations of major cities, many schools—and thus people—are

[6]The PUR data are publicly available on the California DPR's website.

[7]The PUR also includes crop type, area planted, and area treated when the user applies the pesticide to a crop.

[8]Figure C.2 repeats this exercise for the number of pesticide applications by day of week—the trends are quite similar.

[9]Instead of an actual logarithmic transformation, I use the inverse hyperbolic sine so as to include sections with exactly zero pounds of pesticide application.

[10]Using the database of school locations provided by CSCD.

located on the edges of these cities. Many rural schools are located in sections with high levels of pesticide applications.

3.2.3 Wind

The wind data for this project come from NASA's North American Land Data Assimilation Systems, version 2 (NLDAS-2). The NLDAS-2 joins observation-based and model-reanalysis data to generate land-surface models. Part of this process involves generating (forcing) wind-vector data. The resulting wind data are available for each hour since January 1979 at a 1/8$^{\text{th}}$-degree grid covering all of North America (Xia et al., NCEP/EMC(2009)). Specifically, the NLDAS-2 generates two wind vectors—a zonal vector U component (the westerly component) and the meridional V component (the southerly component) (Xia et al., NCEP/EMC(2009)). Jointly the two components determine the wind direction (degrees) and speed (meters per second). The data appendix contains more information on these trigonometric calculations.

3.3 Empirical strategy

As highlighted above, this paper seeks to answer whether there is significant evidence that aerially applied pesticides aggregate and drift in ways that contribute to the poor air quality observed in the southern counties of California's San Joaquin Vally. In order to detect variation in *air quality*, I specifically consider local PM2.5 concentrations at EPA monitors in the study counties (depicted in Figure 3.1). Accordingly, one might model the PM2.5 concentration on a given day t at a given monitor i by

$$\text{Concentration}_{i,t}^{\text{PM2.5}} = f(d_{i,t}(\text{Pesticides}_{i,t}), \text{Weather}_{i,t}) + \varepsilon_{i,t} \tag{3.1}$$

where $d_{i,t}(\text{Pesticides}_{i,t})$ defines an arbitrary distance-based aggregator of the pesticides applied on day t,[11] $\text{Weather}_{i,t}$ refers to the weather near monitor i on day t, f represents an arbitrary function of the pesticides applied (including their distances) and weather on day t relative to monitor i, and $\varepsilon_{i,t}$ catches stochastic variation in PM2.5 concentrations.

3.3.1 Fixed effects

To place some structure on equation 3.1, I allow quantities of pesticides applied at similar distances from a monitor to similarly affect that PM2.5 concentrations at the monitor. Such an assumption effectively creates buffers—or concentric rings—around each monitor, where each ton of pesticide between rings similarly affects PM2.5 concentrations. This design is sometimes referred to as a *doughnut* design. Figure 3.8 illustrates this design for three buffers: (**1**) within 1.5km, (**2**) between 1.5km and 3km, and (**3**) between 3km and 25km.

[11] While likely goes to zero at some distance, effectively creating a buffer around monitor i.

A fixed-effect based estimating equation for this design is

$$
\begin{aligned}
\text{Concentration}_{i,t}^{\text{PM2.5}} = & \beta_1 \sum_{k(t)} \mathbb{1}\left\{\mathscr{D}_{i,k(t)} < 1.5\text{km}\right\} \times p_{k(t)} + \\
& \beta_2 \sum_{k(t)} \mathbb{1}\left\{1.5\text{km} \leq \mathscr{D}_{i,k(t)} < 3\text{km}\right\} \times p_{k(t)} + \\
& \beta_3 \sum_{k(t)} \mathbb{1}\left\{3\text{km} \leq \mathscr{D}_{i,k(t)} < 25\text{km}\right\} \times p_{k(t)} + \\
& \gamma_i + \delta_t + \varepsilon_{i,t}
\end{aligned} \tag{3.2}
$$

where i indexes EPA monitors; $k(t)$ references the k^{th} pesticide application on day t; and $p_{k(t)}$ records the amount of pesticides applied in application $k(t)$. In addition, $\mathscr{D}_{i,k(t)}$ gives the distance (in kilometers) between EPA monitor i and pesticide application $k(t)$ at time t; $\mathbb{1}\{\texttt{foo}\}$ denotes an indicator function for whether `foo` is true; and γ_i and δ_t refer to individual-monitor and temporal fixed effects, respectively. I vary the type of temporal fixed effect across several specifications—ranging from day-of-sample to month-of-year, year, and day-of-week.[12]

The temporal fixed-effect specifications vary the identifying variation for the parameters of interest—the β_j—in equation 3.2. For instance, month-of-year fixed effects control for the average PM2.5 concentrations and pesticide applications for a day in the given calendar month throughout the sample period (conditional on the individual-monitor fixed effects). Thus, identification of the β_j results from (daily) deviations from these observed means. This *doughnut* empirical design lends an additional source of identifying variation—the concentric circles allow an additional ton of pesticides applied in near proximity to the monitor to have a different effect than an additional ton of pesticides applied farther from the monitor. Put simply: identification in this fixed-effects doughnut design results from the question: On days when pesticide applications near the monitor exceed the monthly norm, do we also see PM2.5 concentrations exceed their monthly norms? Moreover, this design offers a natural check for the plausibility of the results: the estimates for β_1, β_2, and β_3 in equation 3.2 should be monotonically decreasing, as aerially applied pesticides diffuse through space—far-away applications should have smaller effects. The results are consistent with this plausibility check. Finally, if pesticides negatively affect (local) air quality, then β_1 should be significantly greater than zero. If $\beta_2 > 0$ or $\beta_3 > 0$, then the estimates imply substantial aerial pesticide drift (movement over space).

Although the fixed-effect identification strategy potentially isolates exogenous variation, it is still susceptible to bias from omitted variables. The concern is that there may be a daily-varying factor excluded from the model that causally affects both PM2.5 concentrations and pesticide applications. If such a variable exists, then the estimates for the β_j may be positively or negatively biased, depending upon the relationships between the omitted variable, PM2.5 concentrations, and

[12] I use *month of year* to refer to the calendar months (*e.g.*, January) and *month of sample* to reference specific month-year combinations (*e.g.*, January 2010).

pesticide applications. The concentric-circles design, in conjunction with day-of-sample fixed effects, may alleviate some omitted-variable bias concerns, as the identifying variation comes both from the amount of pesticide application *and* the distances between the pesticide applications and the sensor. Thus, the omitted variable would need to increase the number of pesticide applications *near the EPA monitor* on the same day it increased the PM2.5 concentration *near the EPA monitor*. Consequently, in addition to demonstrating general robustness, the different fixed-effects specifications that I present in the results—and their different sources of identifying variation—suggest that the results in this paper are not driven by omitted-variable bias.

3.3.2 Wind-angle variation

To further the plausibility of the results, I present a second identification strategy, which extends the fixed-effects *doughnut design* discussed above. This second design further isolates plausibly exogenous identifying variation by incorporating daily deviations from the prevailing wind pattern (angle) at the EPA sensor—separating *upwind*, *downwind*, and *orthogonal wind* pesticide applications.

In order to isolate upwind pesticide applications, I calculate (**a**) the angle between each EPA monitor and each pesticide application[13] and (**b**) the angle of the wind at each EPA monitor.[14] The difference between these two angles gives a measure of whether pesticide application $p_{k(t)}$ occurred upwind of monitor i (in degrees). To integrate this upwind measure into an empirical model, I bin applications into three broad groups: (**1**) **upwind application** where the absolute difference between the wind's angle and the monitor-to-application angle angle is less than 60 degrees; (**2**) **orthogonal application** where the absolute difference between the two angles is between 60 and 120 degrees; and (**3**) **downwind application** where the absolute difference between the two angles is between 120 degrees and 180 degrees.[15] Figure 3.9a illustrates this wind-angle grouping.

Combining this wind-direction information/variation with the fixed-effects doughnut model in equation 3.2, the estimating equation for this new model is

$$
\begin{aligned}
\text{Concentration}_{i,t}^{\text{PM2.5}} = \; & \\
& \alpha_{11} \sum_{k(t)} \mathbb{1}\Big\{\Big(\mathscr{D}_{i,k(t)} < 1.5\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in [0,\ 60]\Big)\Big\} \times p_{k(t)} + \\
& \alpha_{12} \sum_{k(t)} \mathbb{1}\Big\{\Big(\mathscr{D}_{i,k(t)} < 1.5\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in (60,\ 120]\Big)\Big\} \times p_{k(t)} + \\
& \alpha_{13} \sum_{k(t)} \mathbb{1}\Big\{\Big(\mathscr{D}_{i,k(t)} < 1.5\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in (120,\ 180]\Big)\Big\} \times p_{k(t)} +
\end{aligned} \tag{3.3}
$$

[13]Because the PUR data only identify applications at the section level, I use the geographic coordinates of the section's centroid for each application within the section.

[14]The wind's vector points upwind: the angle between a ray pointing toward the wind's origin and due North.

[15]An alternative way to think about measure is the angle between two vectors: (**1**) the vector between the pesticide application and the EPA monitor, and (**2**) the wind vector at the EPA monitor.

$$
\begin{aligned}
&\alpha_{21} \sum_{k(t)} \mathbb{1}\Big\{\Big(1.5\text{km} \le \mathscr{D}_{i,k(t)} < 3\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in [0,\ 60]\Big)\Big\} \times p_{k(t)} + \\
&\alpha_{22} \sum_{k(t)} \mathbb{1}\Big\{\Big(1.5\text{km} \le \mathscr{D}_{i,k(t)} < 3\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in (60,\ 120]\Big)\Big\} \times p_{k(t)} + \\
&\alpha_{23} \sum_{k(t)} \mathbb{1}\Big\{\Big(1.5\text{km} \le \mathscr{D}_{i,k(t)} < 3\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in (120,\ 180]\Big)\Big\} \times p_{k(t)} + \\
&\alpha_{31} \sum_{k(t)} \mathbb{1}\Big\{\Big(3\text{km} \le \mathscr{D}_{i,k(t)} < 25\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in [0,\ 60]\Big)\Big\} \times p_{k(t)} + \\
&\alpha_{32} \sum_{k(t)} \mathbb{1}\Big\{\Big(3\text{km} \le \mathscr{D}_{i,k(t)} < 25\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in (60,\ 120]\Big)\Big\} \times p_{k(t)} + \\
&\alpha_{33} \sum_{k(t)} \mathbb{1}\Big\{\Big(3\text{km} \le \mathscr{D}_{i,k(t)} < 25\text{km}\Big) \wedge \Big(\big|\theta_{i,k(t)}\big| \in (120,\ 180]\Big)\Big\} \times p_{k(t)} + \\
&\gamma_i + \delta_t + \varepsilon_{i,t}
\end{aligned}
$$

where all quantities maintain the same definitions as in equation 3.2, and θ_{it} denotes the difference in the wind angle on day t and monitor i and the pesticide angle between monitor i and pesticide application $k(t)$—as defined and discussed directly above. Put simply, equation 3.3 allows the effect of a pesticide application on PM2.5 concentration to vary by the application's distance from the monitor *and* by the application's degree of *upwind-ness*—again controlling for a variety of individual (γ_i) and temporal (δ_t) fixed effects.

By adding wind-induced variation in the (conditional) amount of pesticides applied, equation 3.3 further relaxes the assumptions required to identify its parameters of interest (the α_j). In order for an omitted variable to bias the ordinary least squares (OLS) estimates of the α_j, there would need to be an observed variable correlated with both (**1**) PM2.5 concentrations and (**2**) the amount of pesticides applied (**3**) upwind of EPA monitors—and with (**4**) the distances between the pesticide applications and the EPA monitors. In the absence of such a process, OLS will provide causally valid, consistent estimates for the extent to which aerial pesticide applications and their drift affect local PM2.5 concentrations.

The two panels of Figure 3.9 depict this wind-induced variation design—illustrating how the design estimates a coefficient for each 60-degree segment of the three distance-based radial groups.[16]

3.3.3 Measurement error

Having outlined this paper's two identification strategies and its datasets, I now discuss an empirically relevant data issue before presenting the results from estimating equations 3.2 and 3.3 via OLS.

Measurement error presents problems for both empirical designs in this paper, but it is particularly important for the design using wind-induced variation. One

[16] I enforce a requirement for symmetry, *e.g.*, pesticide applications within 1.5km have the same effect regardless of whether they occur at -45 degrees or 45 degrees of the wind.

of the main sources of measurement error comes from the lack of spatial precision in the PUR data. As discussed above, the PURs only identify pesticide applications down to the section level. Sections can be as large as 1 mile by 1 mile, meaning geography-based variables are likely to contain substantial noise. The variables of interest in equations 3.2 and 3.3 are both based upon the geographic coordinates of the pesticide applications. For the distance-based indicators in the two regressions, this measurement error simply adds noise to the coordinates—akin to rounding, in some sense. Consequently, measurement error for the fixed-effects doughnut design is classical in nature and will simply attenuate OLS parameter estimates (Woolridge 2010).

For the the angle-based measurements, this geographic measurement error induces classical measurement error *for pesticides applications within the same distance of an EPA monitor.* For pesticide applications closer to the monitors, the attenuation bias will be larger. To see this point, consider three cases. First, if a pesticide application occurs in the same section as an EPA monitor, the uncertainty surrounding the application's location prevents one from knowing whether the application is upwind, downwind, or orthogonal to the monitor. Second, if the application occurs in the section next to the monitor's section, it is possible to bound the angle between the EPA monitor and the application between −90 and 90 degrees—potentially ruling out one of the three upwind categories—but we are still left with substantial noise. Finally, for a pesticide application far from the monitor, there is little uncertainty in the angle between the monitor and application. The result of this class of semi-classical measurement error is that estimates for the parameters α_{1j} will be more attenuated than the estimates for α_{2k}.[17]

In addition, there are several other pertinent sources of noise in the data and, therefore, in the empirical strategies. First, the reanalysis wind data are not actual historical records—adding noise in addition to identifying variation. Second, the data on pesticide applications come from self-reported pesticide-use reports. Self-reported data often contain a degree of noise—and may contain some bias where incentives lead to dishonest reporting. Third, aggregating pesticide applications into distance- and/or angle-based groups potentially leads to a sort of aggregation bias—averaging across the heterogeneous treatment effects within each group. As a result of these channels of noise and attenuation, the parameter estimates in the next section should be taken as lower bounds of the actual effects of aerial pesticide applications on PM2.5 levels.

3.4 Results

Having described the identification strategies of two models—the fixed-effects doughnut design and the wind-angle design—I now present the OLS estimates for equations 3.2 and 3.3.

[17] It is also worth noting that statistical power follows a similar trend due to the close applications covering smaller areas of land than the farther out application—resulting in more observations and greater variation in distance-based groups that are farther from the monitors.

3.4.1 Fixed effects

Table 3.2 presents the results for estimating equation 3.2 with OLS. Specifically, Table 3.2 identifies the effect of an additional ton of aerially applied pesticides—applied in one of three distance-based groups—on the mean PM2.5 concentration at monitor i on day t. Each of the five columns of Table 3.2 represents a different regression from a different fixed-effect specification for equation 3.2. Each specification includes a monitor fixed effect. Columns (**1**), (**2**), and (**3**) include day-of-sample, week-of-sample, and month-of-sample fixed effects, respectively. Column (**4**) incorporates three sets of time-based fixed effects: week-of-year, day-of-week, and year fixed effects. Column (**5**) uses month-of-sample, day-of-week, and year fixed effects.

Because aerial pesticide applications occur at a specific time during the day, one might expect the daily mean to underestimate the effect of pesticide applications on local air quality—averaging across affected and unaffected times of the day and generating a sort of aggregation or attenuation bias. To potentially remedy this problem, Table 3.3 replicates Table 3.2 but uses the daily *maximum* observed PM2.5 concentration rather than the daily *mean*. The problem with this potential remedy is that aerial applications may not move the maximum PM2.5 if they occur at lower points of the PM2.5 diurnal cycle.

In each of the five specifications of Tables 3.2 and 3.3, the point estimate for the distance group nearest to the EPA monitor (within 1.5 kilometers) is significantly different from zero. The point estimates range from 0.12 to 0.23 for the daily mean PM2.5 level, and they range from 0.15 to 0.34 for the daily maximum PM2.5 level. The causal interpretation for the first point estimate in column (**1**) of Table 3.2 is that each additional ton of aerially applied within 1.5 kilometers pesticides increases that day's mean PM2.5 level by approximately 0.23 $\mu g/m^3$ [0.09, 0.37]. For column (**1**) of Table 3.3, the interpretation is that each additional ton of pesticides applied aerially within 1.5 kilometers of the EPA monitor increases the daily maximum observed PM2.5 concentration by approximately 0.31 $\mu g/m^3$ [0.12, 0.51].

The two panels of Figure 3.8 illustrate the spatial relationships of the results in conjunction with the fixed-effect doughnut design for the first columns of Tables 3.2 and 3.3.

Across the ten regressions of Tables 3.2 and 3.3, the results present clear and robust[18] evidence that aerial pesticide applications significantly reduce local air quality (increasing PM2.5 concentrations)—for aerial pesticide applications within 1.5 kilometers of the monitor. None of the results in the two tables present statistically significant evidence that aerial applications affect PM2.5 levels beyond 1.5 kilometers.[19]

[18]Appendix tables C.1 and C.2 demonstrate further robustness by replicating Tables 3.2 and 3.3, respectively, with Windsorized PUR application data—diminishing concerns that outliers in the independent variable drive the results. The Data Appendix describes describes this Windsorization in detail.

[19]While the coefficients on the distance groups farther than 1.5 kilometers are not statistically significantly different from zero at conventional levels, the point estimates are consistently positive and also decrease monotonically with distance. Furthermore, I particularly caution the reader from

3.4.2 Wind-angle variation

Tables 3.4 and 3.5 contain the OLS results of estimating the wind-variation based model in equation 3.3. Each column contains results from a separate fixed-effects specification; all five regressions use daily mean PM2.5 concentration as their dependent variable. Relative to the previously presented results/model in Table 3.2: the results in Tables 3.4 and 3.5 introduce wind-based variation in pesticide exposure.

Figure 3.9a illustrates the spatial relationship and implications of the coefficients estimated in Tables 3.4 and 3.5—specifically depicting the results of column (**1**) of Table 3.4. The results portrayed in Figure 3.9a—and across all the columns Tables 3.4 and 3.5—again suggest that aerially applied pesticides within 1.5km increase local PM2.5 concentrations. Further, the wind-angle results suggest—as one would expect from a more physics-based model—that this increase in PM2.5 levels induced by aerial pesticides particularly stems from upwind pesticide applications. Across the five specifications in Tables 3.4 and 3.5, the point estimates for *Upwind* applications within 1.5 kilometers range from 0.24 to 0.29—slightly larger than the non-wind results from Table 3.2. The causal interpretation for these results—using column (**1**) of Table 3.4—is that each additional ton of aerial pesticides applied upwind within 1.5 kilometers increases the daily mean PM2.5 concentration by 0.29 $\mu g/m^3$ [0.05, 0.55].

The point estimates for the effect of upwind aerial pesticide applications within 1.5 kilometers are statistically significant and notably large across all five specifications in Tables 3.4 and 3.5. The point estimates for downwind applications 3–25 kilometers away are also significant across all five specifications, though the magnitude of the point estimates is quite small. No other estimated effect is consistently significant across all five specifications.

Tables 3.6 and 3.7 replicate Tables 3.4 and 3.5 but with daily *maximum* PM2.5 rather than daily *mean* PM2.5 . Figure 3.9b depicts the spatial effects of aerial-pesticide applications, as estimated in column (**1**) of Table 3.6.

Overall, the results for the daily maximum PM2.5 concentration are fairly similar to the results that use the daily mean. The estimated effect of upwind pesticide applications within 1.5 kilometers is highly statistically significant, and the point estimates for this effect are larger than the estimates for the mean, ranging from 0.29 to 0.49 $\mu g/m^3$. The causal interpretation of this effect (for column (**1**) of Table 3.6) is that each additional ton of pesticides aerially applied upwind within 1.5 kilometers increases the daily local maximum PM2.5 concentration by 0.49 $\mu g/m^3$ [0.19, 0.78]. One surprising outcome is the large (in magnitude) and negative coefficient for pesticides applied within 1.5 kilometers orthogonally to the wind. This effect is statistically significant at the 5-percent level in three of the five specifications.

Finally, while this empirical design offers greater potential for isolating exogenous variation by using changes in wind patterns, one should bear in mind that this design is also more prone to bias from measurement error. This susceptibility to measurement-error induced bias originates in the lack of precise spatial information

reading too much into null results here, as attenuation bias is clearly present.

in the PUR data and potential noise in the reanalysis wind data.[20]

3.5 Discussion and conclusion

Across two empirical designs and a many specifications, this paper finds significant evidence that aerially applied pesticides reduce local air quality—increasing PM2.5 concentrations at EPA monitoring sites. Whether the results use distance-based variation or wind-and-distance variation, I find evidence consistent with aerial pesticide applications increasing both daily mean and daily maximum PM2.5 levels. The point estimates suggest that each additional ton of aerial pesticides applied within 1.5 kilometers increases local PM2.5 concentrations by approximately 0.2 to 0.3 $\mu g/m^3$. To put this coefficient in perspective: for days on which at least one aerial pesticide application occurred in a section[21], the 90$^{\text{th}}$ (99$^{\text{th}}$) percentile of the amount of aerial pesticides is 3 tons (18.4 tons).[22] Thus I estimate that applications at the 90$^{\text{th}}$ percentile increase local PM2.5 concentrations by approximately one $\mu g/m^3$, and applications at the 99$^{\text{th}}$ percentile increase PM2.5 levels by approximatley 5 $\mu g/m^3$—nearly half of the national standard for annual PM2.5 concentrations[23]. Accordingly, while most aerial applications do not appear to substantially increase local PM2.5 concentrations, a small percentage of large applications present cases for concern with regards to degraded air quality. This result is consistent with recent work tying large, local applications of pesticides to adverse health effect: Larsen, Gaines, and Deschênes only find evidence the top five percentiles of pesticide exposure increased adverse birth outcomes.

The results of this paper present some good news and some bad news for public-health advocates. The good news: The results suggest that *most* aerial pesticide applications do not meaningfully increase PM2.5 exposure—and the PM2.5 drift may stay within a radius of approximately 1.5 kilometers. The bad news: The top five-to-ten percent of applications substantially increase local PM2.5 concentrations in a region that already suffers from high levels of exposure to PM2.5 and other pollutants.

That said: attenuation matters. As discussed in the empirical strategy and results sections, there are many reasons to believe that measurement error attenuates the results in this paper. Thus, while this paper finds statistically significant evidence of PM2.5 increases due to aerial pesticide applications, the point estimates are most likely lower bounds for the true effect of aerial pesticide applications on local PM2.5 levels. Therefore, much of the "good news" discussed above should be taken with some caution. In addition, many of the results in this paper rely upon daily

[20]See the *Measurement error* subsection of the Empirical strategy section for a detailed description of these measurement-error issues.

[21]Fifteen percent of section-days have at least one aerial pesticide application.

[22]Between 2000 and 2015, in the average year, the agricultural sector in the five study counties aerially applied over 40,000 tons of pesticides.

[23]The EPA reduced the standard for PM2.5 in late 2012 from 15 $\mu g/m^3$ to 12 $\mu g/m^3$. The (U.S. E.P.A. 2013).

mean PM2.5 , which averages across the 24 hourly PM2.5 readings. While the daily mean PM2.5 level is relevant for regulation and—to some degree—health, contemporaneous exposure to PM2.5 also matters. The results of this paper likely substantially underestimate the effect of aerial pesticides on contemporaneous air quality. For instance, if a user applies pesticides at 8:00 AM—as many of the PURs report—the daily mean will include eight *unaffected* hours preceding the application, followed by 16 *affected* hours—suggesting the true effect may be 50 percent larger that the estimated effect.[24]

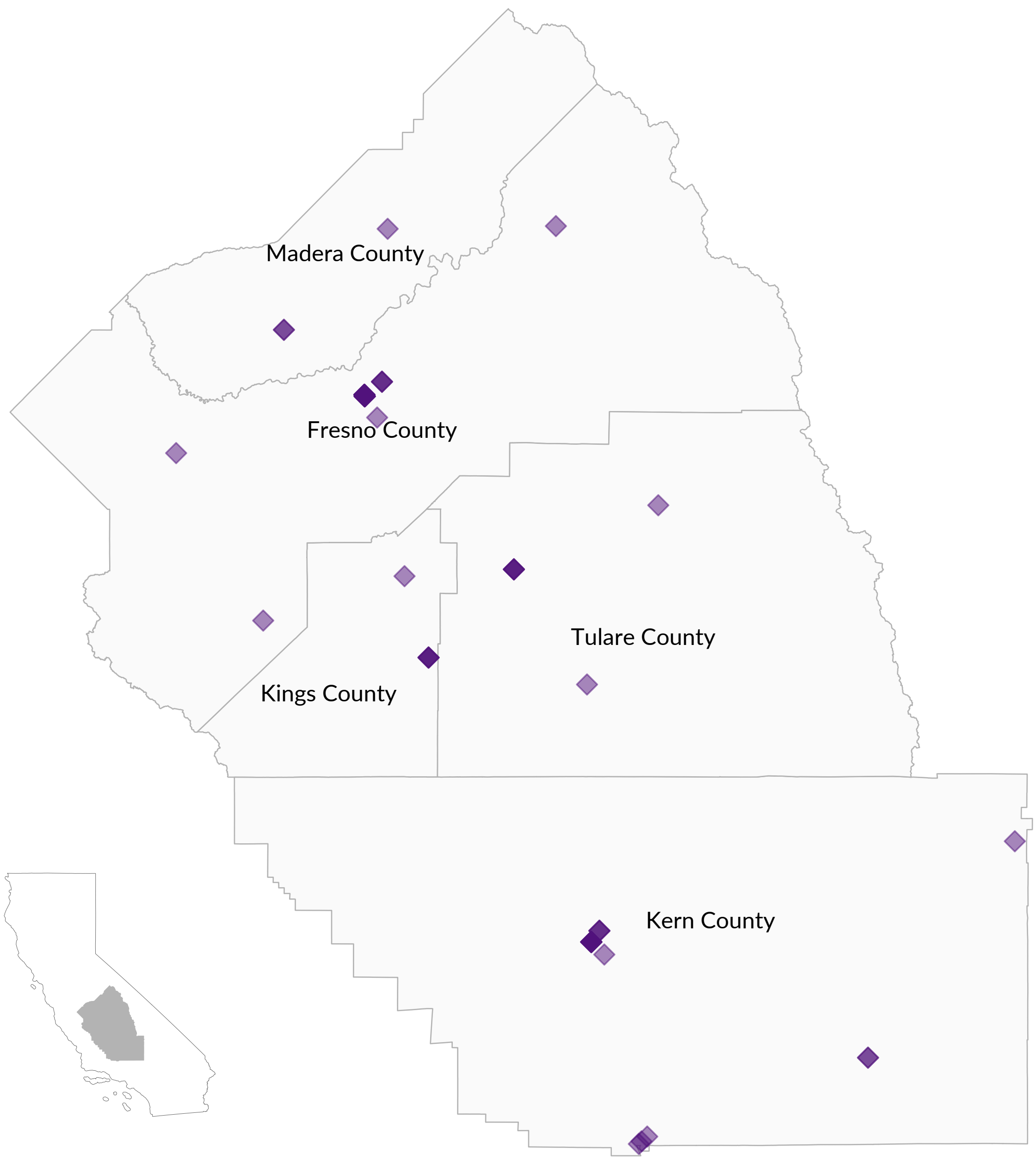
A final *caveat*: PM2.5 is only one measure of local air conditions. For many emissions—and particularly for pesticides—exposures to specific chemicals are of great importance.

Overall, the results in this paper suggest that large aerial pesticide applications substantially increase local PM2.5 exposure—particularly downwind of the application. While policies that target individual pesticide applications or individual users may assist in limiting harmful levels of PM2.5 exposure, they likely miss important effects from the spatiotemporal aggregation of pesticides. This outcome is particularly important for instances where users apply pesticides in close spatial and temporal proximities—especially when these applications are near sensitive/vulnerable individuals. Such areas (conceivably illustrated in Figure 3.6) provide particularly fruitful settings for attention from both policymakers and researchers—offering the potential for investigating the degree to which these observed PM2.5 increases affect health and/or considering efficient spatiotemporal regulations that internalize the costs of pesticide agglomeration and drift.

[24]This calculation assumes a homogeneous effect on the 16 hours following the application. If only a few of the 16 hours are actually affected, the true contemporaneous effect of aerial pesticides applications on local air quality will be many times larger than the point estimates in this paper, *e.g.*, twelve times that of the mean effect if only two hours are *affected*.

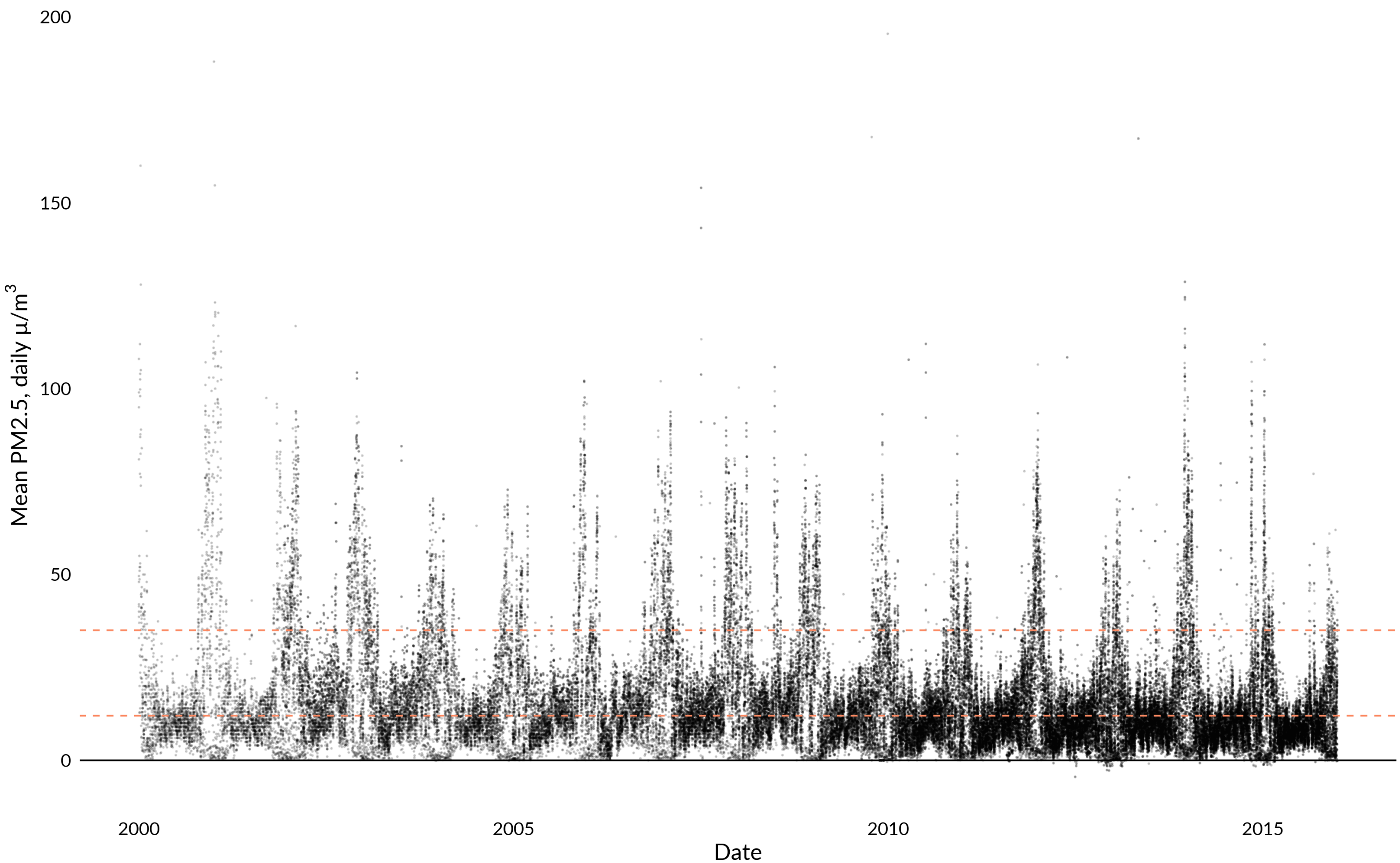
3.6 Figures

Figure 3.1: **EPA PM2.5 monitor locations:** Unique monitors, 2000–2015



Notes: This figure maps the locations of the EPA's PM2.5 monitors (shaded diamonds) in the five counties of the southern San Joaquin Valley. Darker shading denotes the presence of multiple monitors at/near the same site.

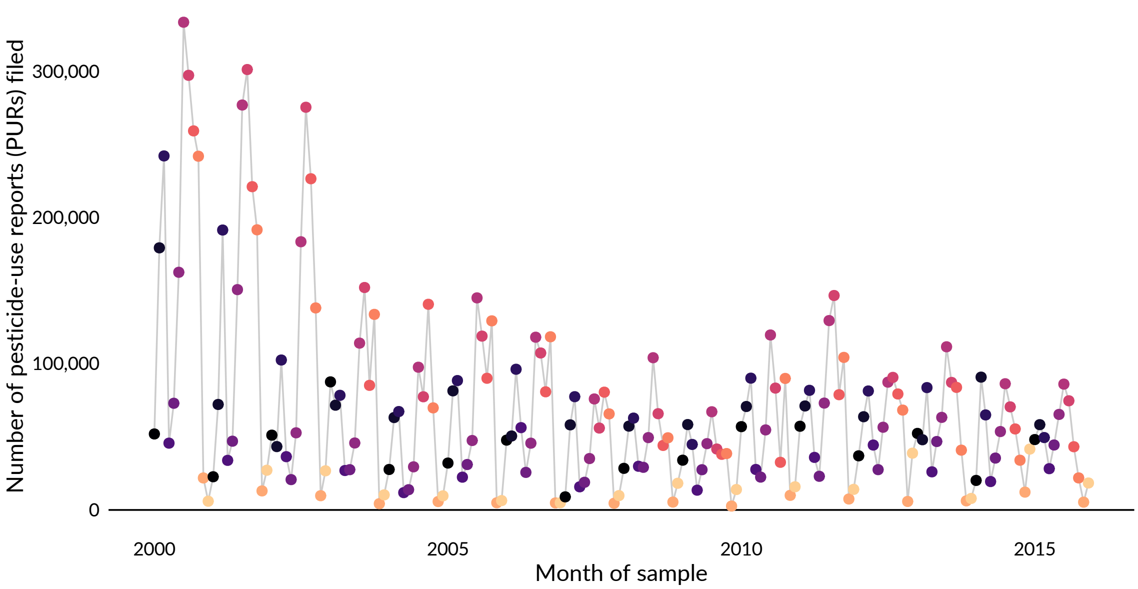
Figure 3.2: **EPA PM2.5 records:** Daily mean PM2.5 concentration at each monitor, 2000–2015



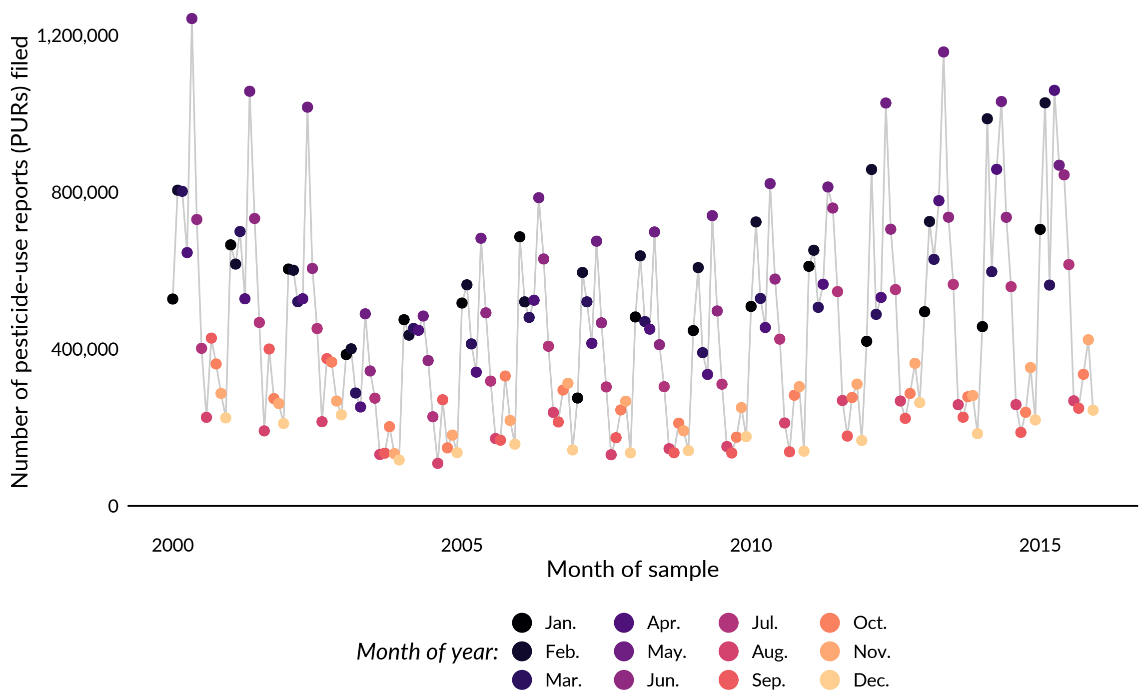
Notes: Each point in this figure represents the mean PM2.5 on the given day (x axis) for a specific monitor. The two dashed horizontal lines denote two different primary National Ambient Air Quality Standards (NAAQS), established January 15, 2013. The lower line establishes the standard (12.0 $\mu g/m^3$) for the the 3-year arithmetic mean. The higher line marks the standard (35 $\mu g/m^3$) for the 3-year mean of the 98$^{\text{th}}$ percentile (U.S. E.P.A. 2013).

Figure 3.3: **Number of pesticide applications:** 2000–2015, by month

(a) **Aerially applied pesticides**



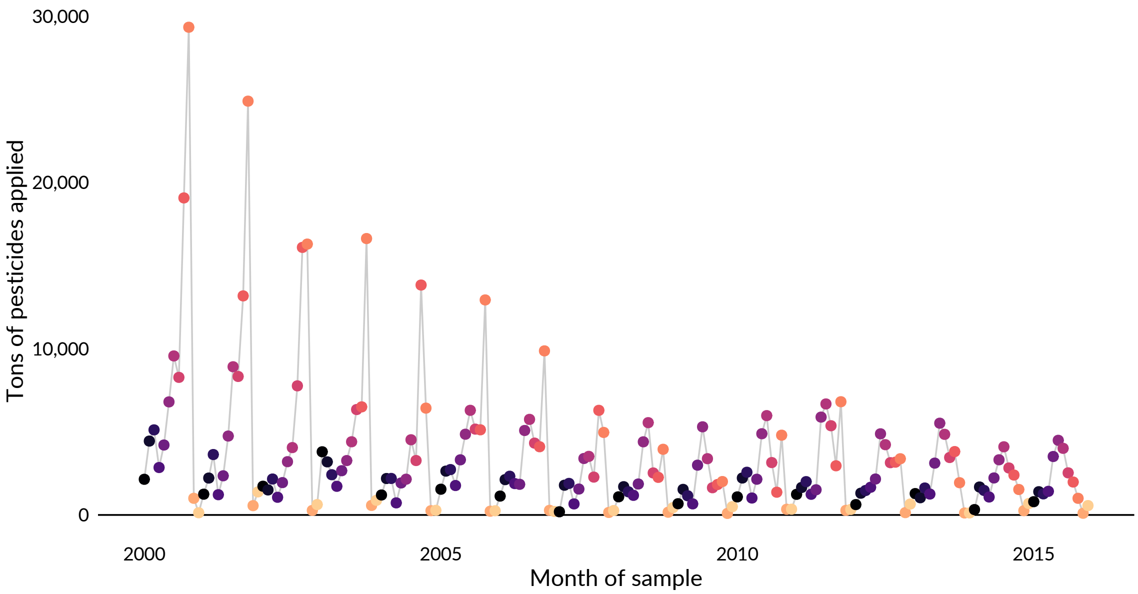
(b) **Ground-applied pesticides**



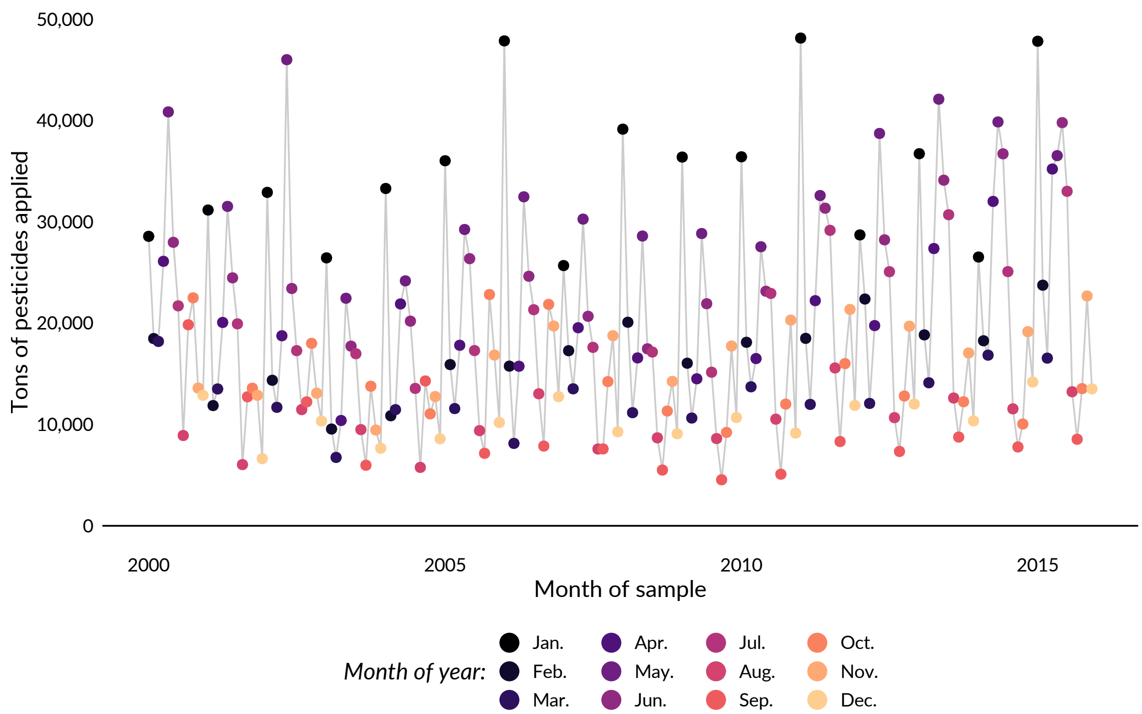
Notes: These figures depict the number of PURs file each month for aerial applications (top) and ground applications (bottom). The counts only include the PURs from the five counties of southern San Joaquin Valley that this paper studies. *Source:* Author using data from California D.P.R. 2013.

Figure 3.4: **Tons of pesticides applied:** 2000–2015, by month

(a) **Aerially applied pesticides**



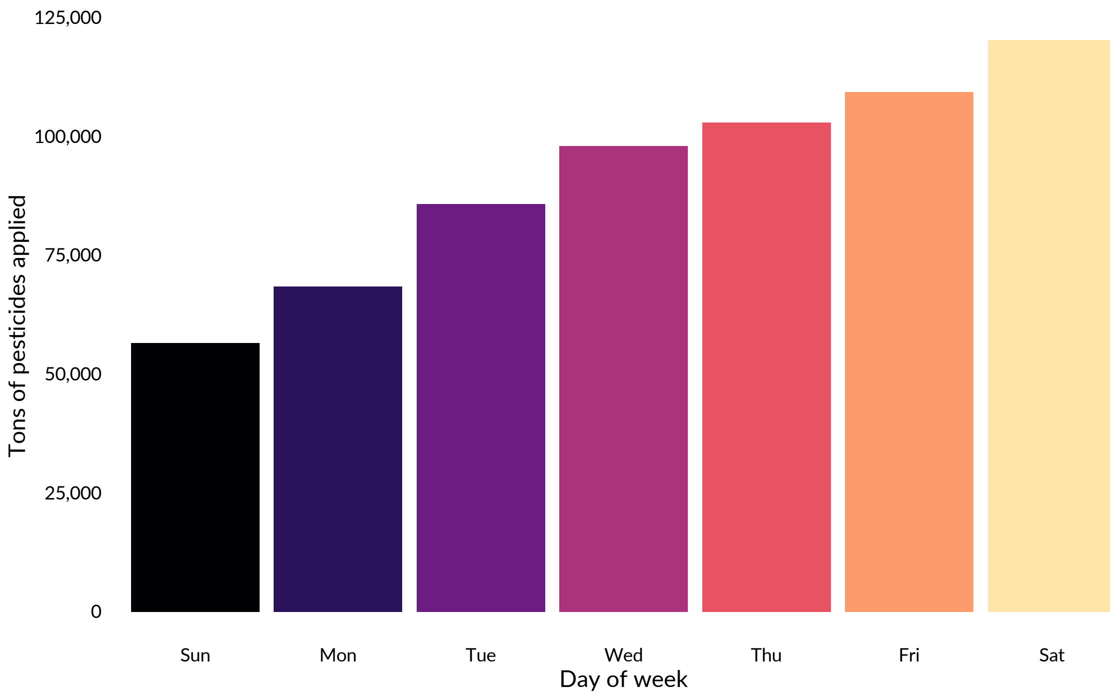
(b) **Ground-applied pesticides**



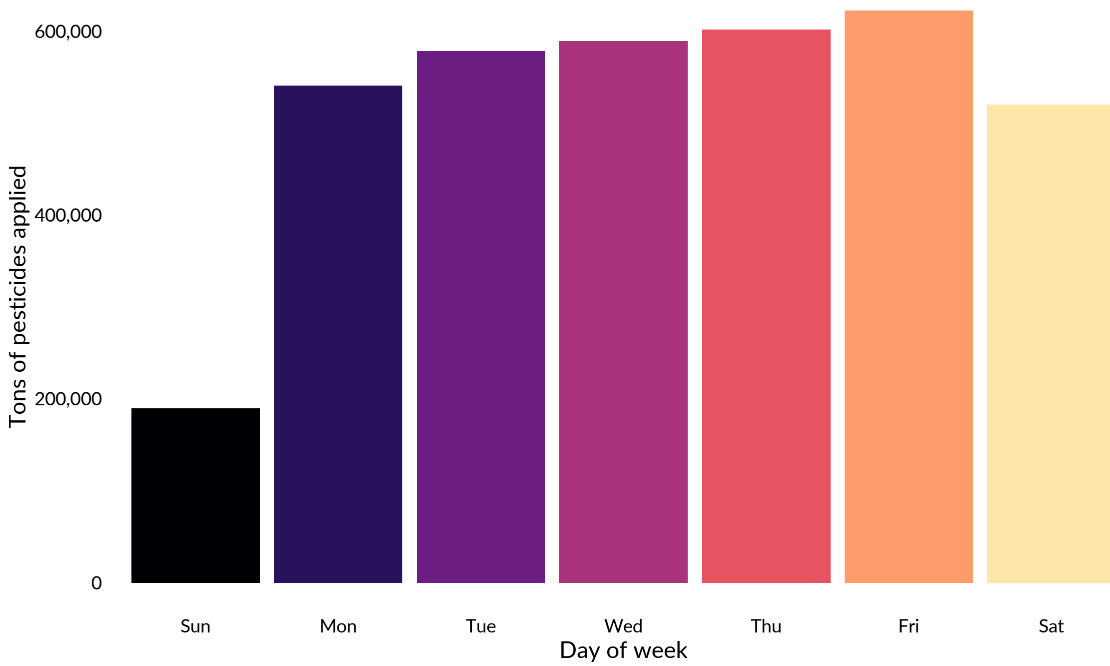
Notes: These figures depict the amount of pesticides (tons) applied each month in the five study counties between 2000 and 2015, separating the pesticide applications by aerial applications (top) and ground applications (bottom).*Source:* Author using data from California D.P.R. 2013.

Figure 3.5: **Tons of pesticides applied:** 2000–2015, by day of week

(a) **Aerially applied pesticides**

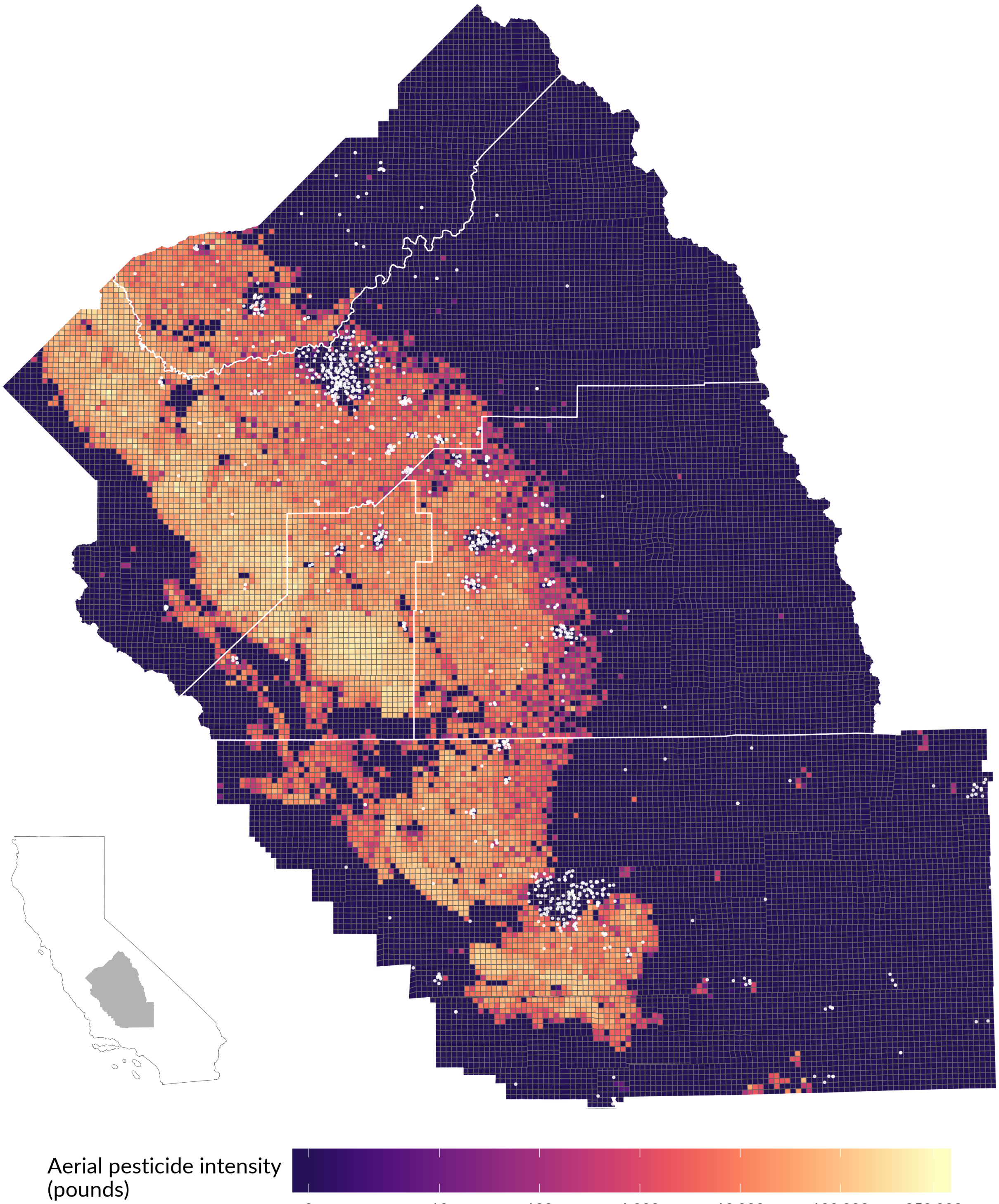


(b) **Ground-applied pesticides**



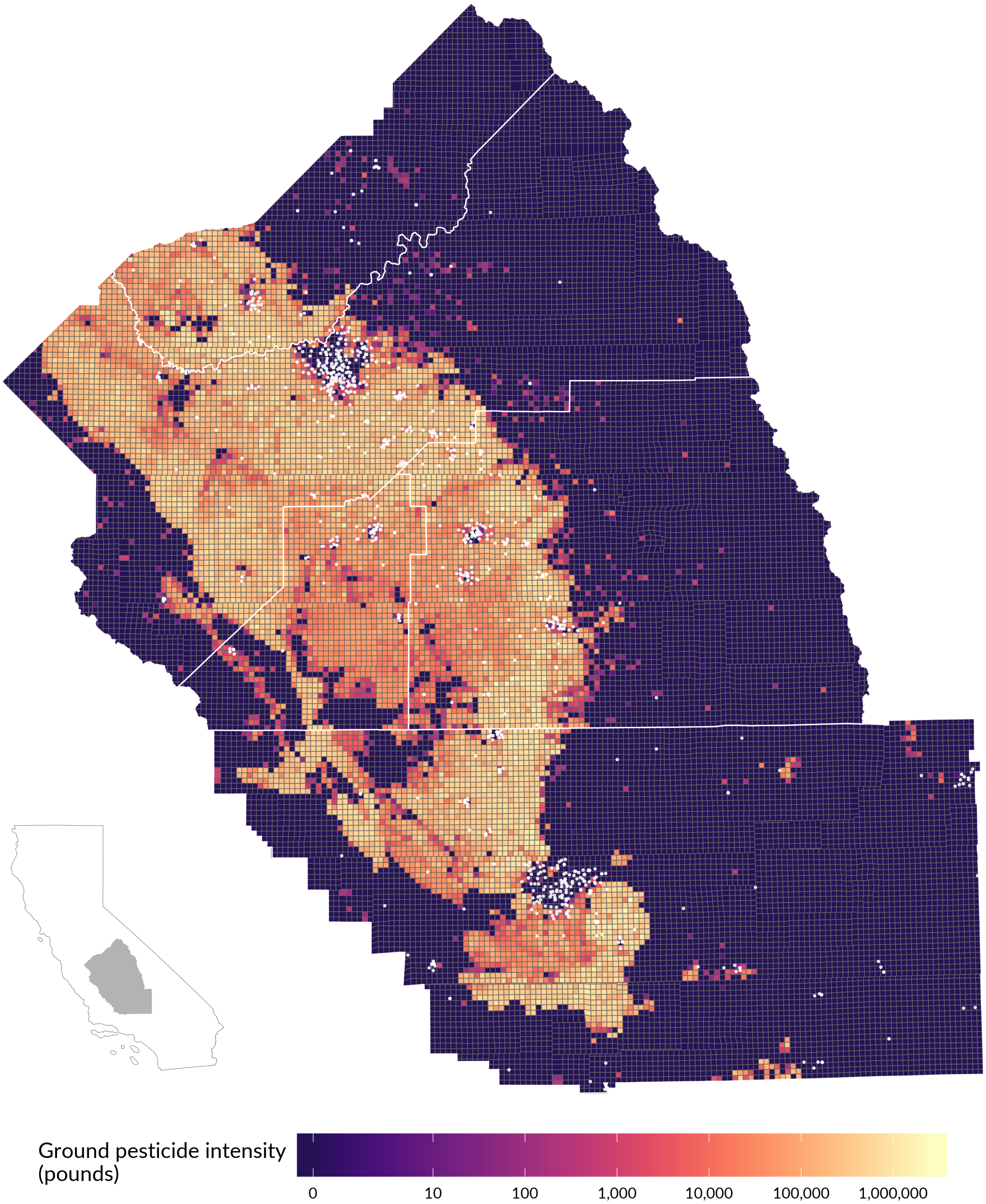
Notes: These figures depict the amount of pesticides (tons) applied by day of week in the five study counties between 2000 and 2015, separating the pesticide applications by aerial applications (top) and ground applications (bottom).*Source:* Author using data from California D.P.R. 2013.

Figure 3.6: **Spatial intensity of aerial pesticide use:** Total pounds of aerial pesticide applications in study counties, 2000–2015



Notes: The shading on the map shows the intensity of aerial pesticide applications within each section from 2000–2015. The shading uses an inverse hyperbolic sine scale (approximately log). White dots denote schools, which I provide as a visual proxy for population. White lines reference county borders.

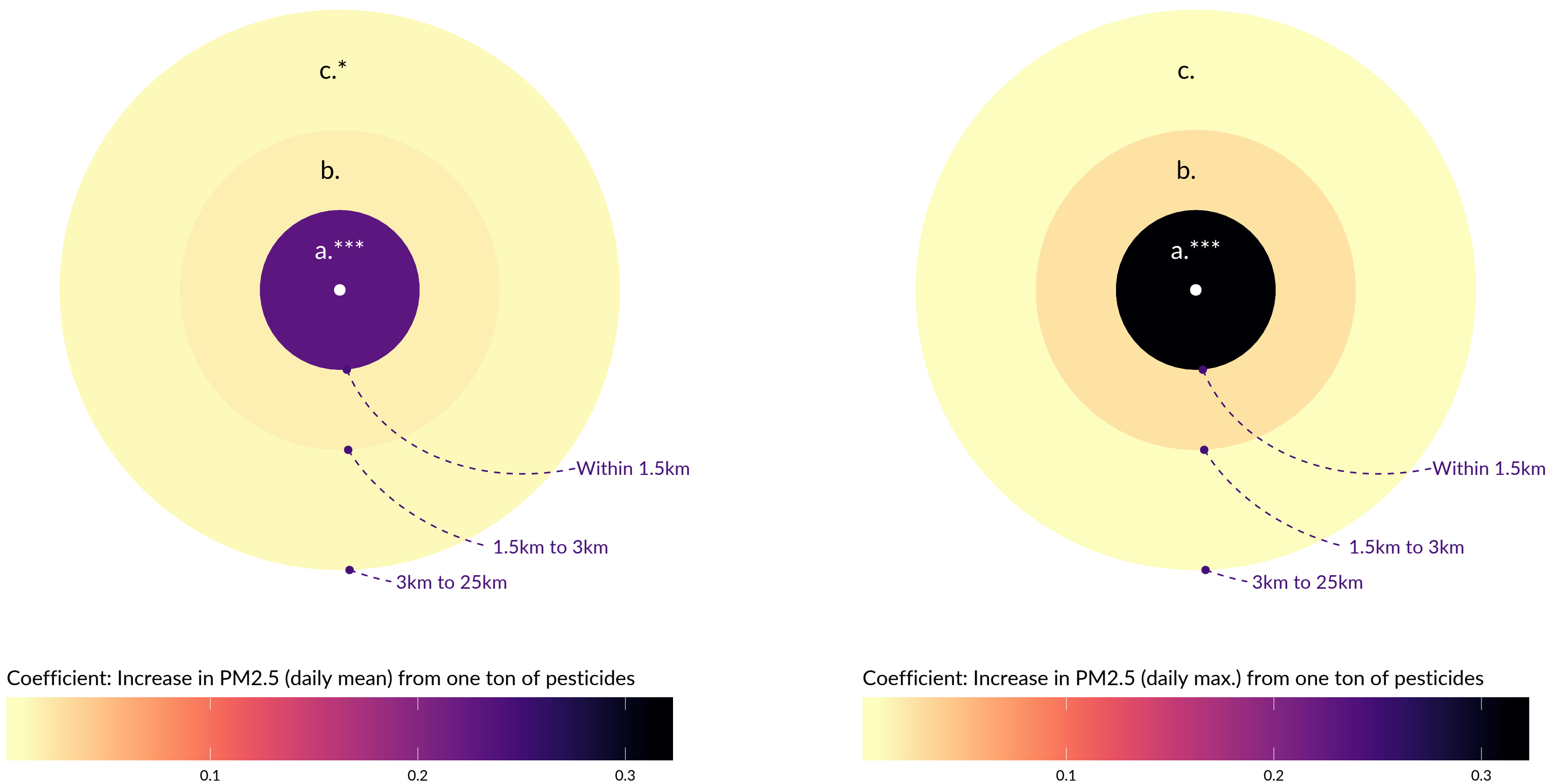
Figure 3.7: **Spatial intensity of ground pesticide use:** Total pounds of ground pesticide applications in study counties, 2000–2015



Notes: The shading on the map shows the intensity of ground pesticide applications within each section from 2000–2015. The shading uses an inverse hyperbolic sine scale (approximately log). White dots denote schools, which I provide as a visual proxy for population. White lines reference county borders.

Figure 3.8: **The effect of pesticides on PM:** Estimated increase in PM2.5 from one ton of aerially applied pesticides using fixed-effects *doughnut* design

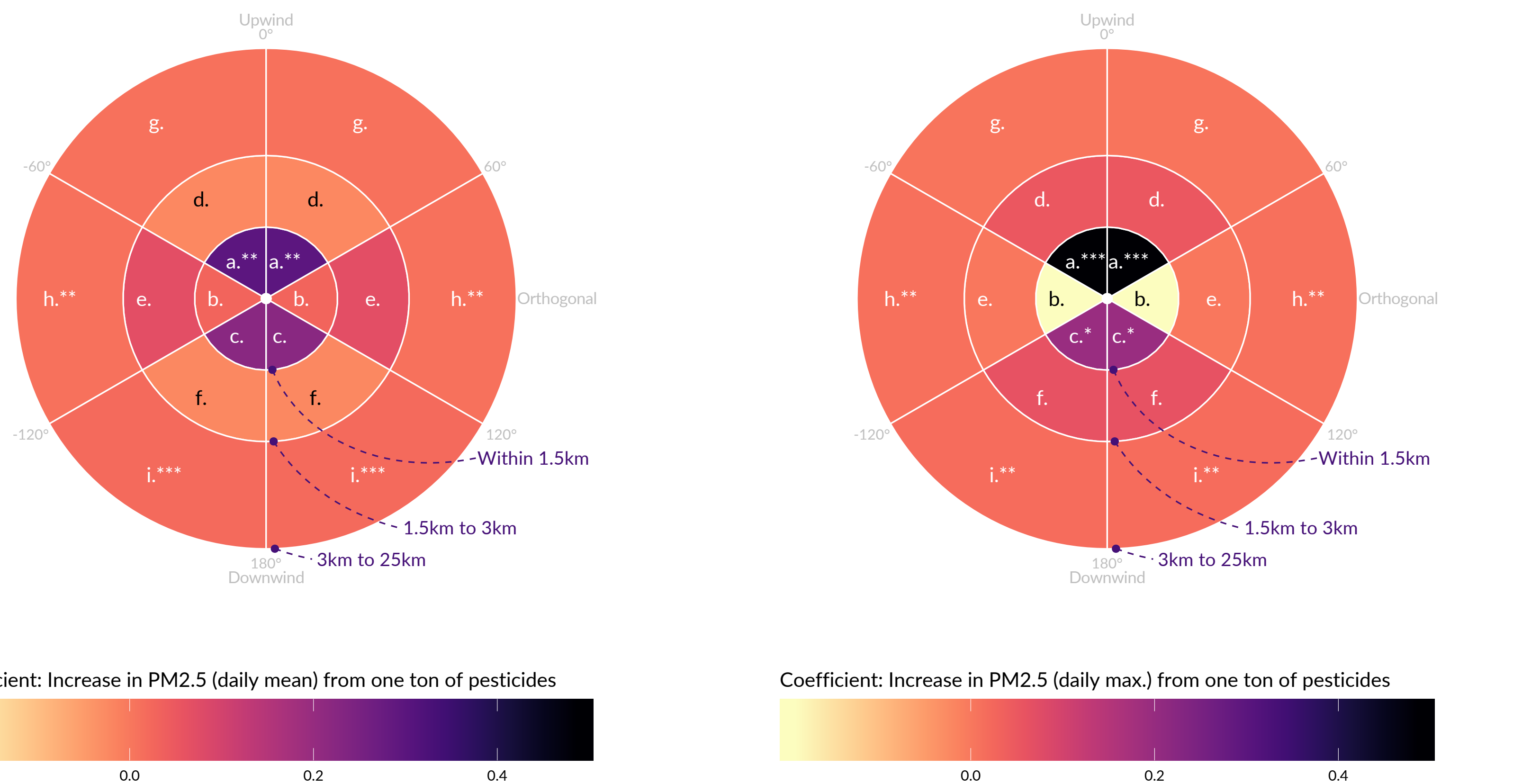
(a) **Daily mean PM2.5** (b) **Daily max. PM2.5**



Notes: This figure illustrates the fixed-effect *doughnut* design of equation 3.2 and the resulting OLS coefficient estimates—as given in column (**1**) of Table 3.2 (a) and column (**1**) Table 3.3 (b).

Figure 3.9: **The effect of pesticides and wind on PM:** Estimated increase in PM2.5 from one ton of aerially applied pesticides using wind variation

(a) **Daily mean PM2.5** (b) **Daily max. PM2.5**



Notes: This figure illustrates the wind-variation design of equation 3.3 and the resulting OLS coefficient estimates—as given in column (**1**) of Table 3.4 (a) and column (**1**) Table 3.6 (b).

3.7 Tables

3.7.1 Descriptive tables

Table 3.1: **Number of monitors and observations** US EPA PM2.5 monitoring network in study counties

Year	N. unique days	N. unique monitors	N. observations
2000	353	23	1,978
2001	361	23	2,261
2002	365	22	4,068
2003	365	22	4,310
2004	366	23	3,596
2005	365	23	4,276
2006	365	23	4,391
2007	365	26	6,483
2008	366	28	6,312
2009	365	30	6,926
2010	365	29	8,916
2011	365	31	10,174
2012	366	36	11,165
2013	365	33	11,337
2014	365	32	10,842
2015	365	30	11,150
All	5,827	53	108,185

Notes: The total number of observations in a year does not equal $N_{\text{days}} \times N_{\text{monitors}}$ because some monitors run every sixth day, rather than every day. For that same reason—and if one or more daily monitors or out of operation—the EPA will observe fewer than 365 days in the year.

3.7.2 OLS results

Table 3.2: **Increases in mean PM2.5**: Same-day, aerially applied pesticides

Dependent variable: Mean daily PM2.5 level	(1)	(2)	(3)	(4)	(5)
Tons of pesticide (**a**)	0.2298***	0.1167**	0.1952**	0.1765**	0.2173**
within 1.5km	(0.07384)	(0.04824)	(0.09388)	(0.07626)	(0.09743)
Tons of pesticide (**b**)	0.0187	0.0290	0.0491	0.0471	0.0428
between 1.5km and 3km	(0.08380)	(0.05007)	(0.05252)	(0.06065)	(0.06687)
Tons of pesticide (**c**)	0.0124*	0.0173**	0.0087	0.0088	0.0055
between 3km and 25km	(0.00734)	(0.00679)	(0.00722)	(0.00733)	(0.00787)
Monitor FE	T	T	T	T	T
Day-of-sample FE	T	F	F	F	F
Week-of-sample FE	F	T	F	F	F
Month-of-sample FE	F	F	T	F	F
Week-of-year FE	F	F	F	T	F
Month-of-year FE	F	F	F	F	T
Day-of-week FE	F	F	F	T	T
Year FE	F	F	F	T	T
N	26,242	26,242	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table 3.3: **Increases in max. PM2.5**: Same-day, aerially applied pesticides

Dependent variable: Maximum daily PM2.5 level

	(1)	(2)	(3)	(4)	(5)
Tons of pesticide (**a**)	0.3149***	0.1321***	0.2104**	0.1607**	0.2036**
within 1.5km	(0.09769)	(0.04612)	(0.08501)	(0.06697)	(0.08453)
Tons of pesticide (**b**)	0.0277	0.0454	0.0713	0.0740	0.0675
between 1.5km and 3km	(0.09168)	(0.05163)	(0.05048)	(0.05977)	(0.06364)
Tons of pesticide (**c**)	0.0099	0.0142*	0.0040	0.0020	−0.0020
between 3km and 25km	(0.00809)	(0.00759)	(0.00851)	(0.00809)	(0.00895)
Monitor FE	T	T	T	T	T
Day-of-sample FE	T	F	F	F	F
Week-of-sample FE	F	T	F	F	F
Month-of-sample FE	F	F	T	F	F
Week-of-year FE	F	F	F	T	F
Month-of-year FE	F	F	F	F	T
Day-of-week FE	F	F	F	T	T
Year FE	F	F	F	T	T
N	26,242	26,242	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table 3.4: **Increases in mean PM2.5 from aerial pesticides**: Wind-variation results

Dependent variable: Mean daily PM2.5 level

	(1)	(2)	(3)
Tons of pesticide (**a**)	0.2983**	0.2380***	0.2781***
within 1.5km; Upwind	(0.12671)	(0.07876)	(0.09443)
Tons of pesticide (**b**)	0.0306	−0.4253***	−0.0473
within 1.5km; Orthogonal	(0.12767)	(0.05577)	(0.11644)
Tons of pesticide (**c**)	0.2227	0.1212*	0.1606
within 1.5km; Downwind	(0.14474)	(0.07032)	(0.19808)
Tons of pesticide (**d**)	−0.0216	0.0349	0.0783
between 1.5km and 3km; Upwind	(0.10813)	(0.10567)	(0.09928)
Tons of pesticide (**e**)	0.0691	0.0202	0.0462
between 1.5km and 3km; Orthogonal	(0.04851)	(0.04303)	(0.03805)
Tons of pesticide (**f**)	−0.0210	0.0464	−0.0059
between 1.5km and 3km; Downwind	(0.25762)	(0.14887)	(0.11126)
Tons of pesticide (**g**)	0.0107	0.0122	0.0018
between 3km and 25km; Upwind	(0.01057)	(0.00839)	(0.01026)
Tons of pesticide (**g**)	0.0105**	0.0200***	0.0136
between 3km and 25km; Orthogonal	(0.00500)	(0.00650)	(0.00846)
Tons of pesticide (**i**)	0.0216***	0.0310***	0.0248***
between 3km and 25km; Downwind	(0.00753)	(0.00889)	(0.00828)
Monitor FE	T	T	T
Day-of-sample FE	T	F	F
Week-of-sample FE	F	T	F
Month-of-sample FE	F	F	T
N	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table 3.5: **Increases in mean PM2.5 from aerial pesticides**: Wind-variation results

Dependent variable: Mean daily PM2.5 level		
	(1)	(2)
Tons of pesticide (**a**) within 1.5km; Upwind	0.2465**	0.2819**
	(0.11788)	(0.12110)
Tons of pesticide (**b**) within 1.5km; Orthogonal	−0.0721	−0.0578
	(0.09426)	(0.10508)
Tons of pesticide (**c**) within 1.5km; Downwind	0.1883	0.2392
	(0.14685)	(0.18599)
Tons of pesticide (**d**) between 1.5km and 3km; Upwind	0.0981	0.0963
	(0.10155)	(0.11903)
Tons of pesticide (**e**) between 1.5km and 3km; Orthogonal	0.0891*	0.0755
	(0.05109)	(0.05202)
Tons of pesticide (**f**) between 1.5km and 3km; Downwind	−0.1957	−0.1793
	(0.17549)	(0.16402)
Tons of pesticide (**g**) between 3km and 25km; Upwind	0.0019	−0.0005
	(0.01037)	(0.01166)
Tons of pesticide (**g**) between 3km and 25km; Orthogonal	0.0135	0.0091
	(0.00896)	(0.00952)
Tons of pesticide (**i**) between 3km and 25km; Downwind	0.0249**	0.0198**
	(0.01032)	(0.00978)
Monitor FE	T	T
Week-of-year FE	T	F
Month-of-year FE	F	T
Day-of-week FE	T	T
Year FE	T	T
N	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.*, '(**a**)') reference labeled areas in the figures associated with these results.

Table 3.6: **Increases in max. PM2.5 from aerial pesticides**: Wind-variation results

Dependent variable: Maximum daily PM2.5 level			
	(1)	(2)	(3)
Tons of pesticide (**a**) within 1.5km; Upwind	0.4869***	0.3073***	0.3458***
	(0.15024)	(0.08086)	(0.10892)
Tons of pesticide (**b**) within 1.5km; Orthogonal	−0.1899	−0.6967***	−0.2682*
	(0.25160)	(0.10545)	(0.13782)
Tons of pesticide (**c**) within 1.5km; Downwind	0.1961*	0.1174	0.1401
	(0.11802)	(0.08083)	(0.19523)
Tons of pesticide (**d**) between 1.5km and 3km; Upwind	0.0508	0.1402	0.1958
	(0.14342)	(0.12992)	(0.12136)
Tons of pesticide (**e**) between 1.5km and 3km; Orthogonal	0.0024	−0.0324	0.0116
	(0.07906)	(0.04590)	(0.04432)
Tons of pesticide (**f**) between 1.5km and 3km; Downwind	0.0627	0.0449	−0.0498
	(0.30296)	(0.28186)	(0.18293)
Tons of pesticide (**g**) between 3km and 25km; Upwind	0.0068	0.0060	−0.0055
	(0.01207)	(0.00989)	(0.01270)
Tons of pesticide (**g**) between 3km and 25km; Orthogonal	0.0122**	0.0224***	0.0133
	(0.00566)	(0.00749)	(0.00973)
Tons of pesticide (**i**) between 3km and 25km; Downwind	0.0171**	0.0303***	0.0220**
	(0.00803)	(0.00997)	(0.00977)
Monitor FE	T	T	T
Day-of-sample FE	T	F	F
Week-of-sample FE	F	T	F
Month-of-sample FE	F	F	T
N	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table 3.7: **Increases in max. PM2.5 from aerial pesticides**: Wind-variation results

Dependent variable: Maximum daily PM2.5 level

	(1)	(2)
Tons of pesticide (**a**)	0.2856***	0.3221***
within 1.5km; Upwind	(0.10680)	(0.11132)
Tons of pesticide (**b**)	−0.3417***	−0.2873**
within 1.5km; Orthogonal	(0.10677)	(0.12585)
Tons of pesticide (**c**)	0.1484	0.1934
within 1.5km; Downwind	(0.13873)	(0.17877)
Tons of pesticide (**d**)	0.2173*	0.2081
between 1.5km and 3km; Upwind	(0.12125)	(0.13675)
Tons of pesticide (**e**)	0.0603	0.0554
between 1.5km and 3km; Orthogonal	(0.06407)	(0.06387)
Tons of pesticide (**f**)	−0.2374	−0.2399
between 1.5km and 3km; Downwind	(0.28403)	(0.25004)
Tons of pesticide (**g**)	−0.0086	−0.0114
between 3km and 25km; Upwind	(0.01273)	(0.01450)
Tons of pesticide (**g**)	0.0130	0.0075
between 3km and 25km; Orthogonal	(0.00953)	(0.00994)
Tons of pesticide (**i**)	0.0208**	0.0149
between 3km and 25km; Downwind	(0.01057)	(0.01019)
Monitor FE	T	T
Week-of-year FE	T	F
Month-of-year FE	F	T
Day-of-week FE	T	T
Year FE	T	T
N	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

B | Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. “Comparative Politics and the Synthetic Control Method”. *American Journal of Political Science* 59 (2): 495–510.

— . 2011. “Synth: An R Package for Synthetic Control Methods in Comparative Case Studies”. *Journal of Statistical Software* 42 (13).

— . 2010. “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program”. *Journal of the American Statistical Association* 105 (490): 493–505.

Abadie, Alberto, and Javier Gardeazabal. 2003. “The Economic Costs of Conflict: A Case Study of the Basque Country”. *American Economic Review* 93 (1): 113–132.

Alberini, Anna, Will Gans, and Daniel Velez-Lopez. 2011. “Residential consumption of gas and electricity in the U.S.: The role of prices and income”. *Energy Economics* 33 (5): 870–881. http://dx.doi.org/10.1016/j.eneco.2011.01.015.

Allcott, Hunt, and Dmitry Taubinsky. 2015. “Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market”. *American Economic Review* 105 (8): 2501–2538.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly harmless econometrics: an empiricist’s companion*. Princeton University Press.

Bai, Jushan. 2009. “Panel Data Models With Interactive Fixed Effects”. *Econometrica* 77 (4): 1229–1279.

Bao, Xuewei, and David W. Eaton. 2016. “Fault activation by hydraulic fracturing in western Canada”. *Science* 354 (6318): 1406–1409.

Becker, Gary S. 1968. “Crime and Punishment: An Economic Approach”. *Journal of Political Economy* 76 (2): 169–217.

Bentham, Jeremy. 1840. *Theory of Legislation*. Weeks, Jordan, & Company.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2003. “How much should we trust differences-in-differences estimates?” *The Quarterly Journal of Economics* 119 (1): 249–275.

Boiteux, M. 1971. “On the management of public monopolies subject to budgetary constraints”. *Journal of Economic Theory* 3 (3): 219–240.

Borenstein, Severin, and Lucas W. Davis. 2012. "The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets". *Journal of Law and Economics* 55.

Brown, SPA, and MK Yücel. 1993. "The Pricing of Natural Gas in U.S. Markets". *Economic Review-Federal Reserve Bank of Dallas*, no. April: 41–52. http://www.dallasfed.org/assets/documents/research/er/1993/er9302c.pdf.

California Department of Pesticide Regulation. 2000. *Pesticide Use Reporting: An Overview of California's Unique Full Reporting System*. Tech. rep. Sacramento, California.

— . 2013. "Pesticide Use Reporting (PUR)". Visited on 02/19/2018. https://www.cdpr.ca.gov/docs/pur/purmain.htm.

California Energy Commission. 2015. "California Building Climate Zone Areas". Accessed March 01, 2017. http://www.energy.ca.gov/maps/renewable/building_climate_zones.html.

— . 2017. "Title 24: Building Energy Efficiency Program". Accessed March 01, 2017. http://www.energy.ca.gov/title24/.

California School Campus Database. 2016. Visited on 04/24/2018. http://www.californiaschoolcampusdatabase.org.

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market". *Industrial and Labor Relations Review* 43 (2): 245–257.

Chen, Yuyu, et al. 2013. "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China 's Huai River Policy". *Proceedings of the National Academy of Sciences of the United States of America* 110 (32): 12936–12941. http://www.pnas.org/content/early/2013/07/03/1300018110.short.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and taxation: Theory and evidence". *American Economic Review* 99 (4): 1145–1177. arXiv: arXiv:1011.1669v3.

City and County of Denver. 2018. "Denver Marijuana Information: Know the Law". Visited on 07/25/2018. https://www.denvergov.org/content/denvergov/en/denver-marijuana-information/know-the-law.html.

Currie, Janet, and W. Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass". *American Economic Journal: Applied Economics* 3 (1): 65–90. doi:10.1257/app.3.1.65.

Davis, Lucas W., and Erich Muehlegger. 2010. "Do Americans consume too little natural gas? An empirical test of marginal cost pricing". *RAND Journal of Economics*.

Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita". *Econometrica* 78 (6): 1863–1903. http://doi.wiley.com/10.3982/ECTA8121.

Environmental Protection Agency. 2013. "Natural Gas Extraction - Hydraulic Fracturing". Accessed August 05, 2016. https://www.epa.gov/hydraulicfracturing.

Esri. 2017. "USA ZIP Code Areas". https://www.arcgis.com/home/item.html?id=8d2012a2016e484dafaac0451f9aea24.

Finlaw, Jack, and Barbara Brohl. 2013. *Task Force Report on the Implementation of Amendment 64: Regulation of Marijuana in Colorado*. State of Colorado.

Glanz, James. 2009. "Quake Fears Stall Energy Extraction Project". *The New York Times* July 13. http://www.nytimes.com/2009/07/14/business/energy-environment/14drill.html.

Gobillon, Laurent, and Thierry Magnac. 2016. "Regional Policy Evaluation: Interactive Fixed Effects and Synthetic Controls". *The Review of Economics and Statistics* 98 (3): 535–551.

Graff Zivin, Joshua, and Matthew Neidell. 2013. "Environment, Health, and Human Capital". *Journal of Economic Literature* 51 (3): 689–730.

Guaita, Rosana, et al. 2011. "Short-term impact of particulate matter (PM2.5) on respiratory mortality in Madrid". *International Journal of Environmental Health Research* 21 (4): 260–274.

Hausman, Catherine, and Ryan Kellogg. 2015. "Welfare and Distributional Implications of Shale Gas". *Brookings Papers on Economic Activity*, no. Spring: 71–125.

Imbens, Guido W, and Jeffrey M Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation". *Journal of Economic Literature* 47 (1): 5–86.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker. 2017. "Every Breath You Take—Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970". *Journal of Political Economy* 125 (3): 848–902.

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing". *American Economic Review* 104 (2): 537–563. http://pubs.aeaweb.org/doi/abs/10.1257/aer.104.2.537.

Kim, Dukpa, and Tatsushi Oka. 2014. "Divorce Law Reforms and Divorce Rates in the USA: An Interactive Fixed Effects Approach". *Journal of Applied Econometrics* 29:231–245.

Kim, Ki Hyun, Ehsanul Kabir, and Shamin Kabir. 2015. "A review on the human health impact of airborne particulate matter". *Environment International* 74:136–143.

Larsen, Ashley E., Steven D. Gaines, and Olivier Deschênes. 2017. "Agricultural pesticide use and adverse birth outcomes in the San Joaquin Valley of California". *Nature Communications* 8 (1): 1–8.

Laumbach, Robert J., and Howard M. Kipen. 2012. "Respiratory health effects of air pollution: Update on biomass smoke and traffic pollution". *Journal of Allergy and Clinical Immunology* 129 (1): 3–11.

Levine, Steven, Paul Carpenter, and Anul Thapa. 2014. *Understanding natural gas markets*. Tech. rep. American Petroleum Institute. http://www.api.org/oil-and-natural-gas/energy-primers/natural-gas-markets.

Lu, Feng, et al. 2015. "Systematic review and meta-analysis of the adverse health effects of ambient PM2.5and PM10pollution in the Chinese population". *Environmental Research* 136:196–204.

Meier, Helena, and Katrin Rehdanz. 2010. "Determinants of residential space heating expenditures in Great Britain". *Energy Economics* 32 (5): 949–959. http://dx.doi.org/10.1016/j.eneco.2009.11.008.

National Academy of Sciences. 2016. "What You Need to Know About Energy: Natural Gas". Accessed August 05, 2016. http://needtoknow.nas.edu/energy/energy-sources/fossil-fuels/natural-gas/.

National Archive of Criminal Justice Data (NACJD). National Incident-Based Reporting System (NIBRS), 2000–2015; Extract Files. 2017. Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research [distributor].

National Organization for the Reform of Marijuana Laws (NORML). 2018. "Colorado Laws & Penalties". Visited on 07/25/2018. http://norml.org/laws/item/colorado-penalties.

Office of Governor John Hickenlooper. 2012a. "Gov. Hickenlooper Signs Amendment 64 Proclamation, Creates Task Force to Recommend Needed Legislative Actions". State of Colorado. Press release accessed through Vote Smart. Visited on 07/25/2018. https://votesmart.org/public-statement/757027.

— . 2013. "Gov. Hickenlooper Signs Numerous Bills". State of Colorado. Press release. Visited on 07/25/2018. https://www.colorado.gov/governor/news/gov-hickenlooper-signs-numerous-bills-4.

Pacific Gas and Electric Company. 2016. "Guide to California Climate Zones". Accessed August 13, 2016. http://www.pge.com/myhome/edusafety/workshopstraining/pec/toolbox/arch/climate/index.shtml.

Pesaran, M. Hashem. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure". *Econometrica* 74 (4): 967–1012.

Pimentel, David, et al. 1992. "Environmental and Economic Costs of Pesticide Use". *BioScience* 42 (10): 750–760.

Pope III, C. Arden. 1989. "Respiratory disease associated with community air pollution and a steel mill, Utah Valley". *American Journal of Public Health* 79 (5): 623–628.

Pope III, C. Arden, et al. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution". *The Journal of the American Medical Association* 287 (9): 1132–1141.

Pot Guide. 2018. "Marijuana Laws in Colorado". Visited on 07/25/2018. https://www.coloradopotguide.com/marijuana-laws-in-colorado/.

PRISM Climate Group. 2004. "PRISM Gridded Climate Data". Oregon State University. http://prism.oregonstate.edu.

Ramsey, F. P. 1927. "A Contribution to the Theory of Taxation". *The Economic Journal* 37 (145): 47–61.

Ransom, Michael R., and C. Arden Pope. 1992. "Elementary school absences and PM10 pollution in Utah Valley". *Environmental Research* 58:204–219.

Rehdanz, Katrin. 2007. "Determinants of residential space heating expenditures in Germany". *Energy Economics* 29 (2): 167–182. http://www.sciencedirect.com/science/article/pii/S0140988306000405.

Sallee, James M. 2013. "Rational Inattention and Energy Efficiency". *Journal of Law and Economics* 57 (3): 781–820.

Seaton, A, et al. 1995. "Particulate air pollution and acute health effects." *Lancet* 345 (8943): 176–178.

Shi, Wei, and Lung-fei Lee. 2017. "The effects of gun control on crimes: a spatial interactive fixed effects approach". *Empirical Economics*: 1–31.

Solon, Garry, Steven J. Haider, and Jeffrey Wooldridge. 2015. "What are we weighting for?" *Journal of Human Resources* 50 (2): 301–316.

Southern California Gas Company. 2016. "California Alternate Rates for Energy (CARE)". Accessed August 08, 2016. https://www.socalgas.com/save-money-and-energy/assistance-programs/california-alternate-rates-for-energy.

State of Colorado. 2012b. *Amendment 64: Use and Regulation of Marijuana*.

Tong, Scott. 2018. "The EPA is at the center of a debate over pesticide regulations". Visited on 07/23/2018. https://www.marketplace.org/2018/04/23/sustainability/epa-center-debate-over-pesticide-regulations.

United States. Environmental Protection Agency. 2016. *Clean Air Act Stationary Source Compliance Monitoring System*. Tech. rep.

— . 2013. *National Ambient Air Quality Standards for Particulate Matter, Final Rule*. Tech. rep. Federal Register, Vol. 78, No. 10 (3086).

U.S. Energy Information Administration. 2016a. "Average annual natural gas spot price in 2015 was at lowest level since 1999". Accessed November 04, 2017. https://www.eia.gov/todayinenergy/detail.php?id=24412#.

— . 2016b. "Electric Power Monthly - Net Generation by Energy Source: Total (All Sectors), 2006-May 2016". Accessed August 05, 2016. https://www.eia.gov/electricity/monthly/epm_table_grapher.cfm?t=epmt_1_1.

— . 2016c. "Frequently Asked Questions". Accessed March 01, 2017. https://www.eia.gov/tools/faqs/faq.cfm?id=45&t=8.

— . 2016d. "Henry Hub Natural Gas Spot Price". Accessed August 10, 2016. https://www.eia.gov/dnav/ng/hist/rngwhhdD.htm.

— . 2017. “Maps: Layer Information for Interactive State Maps - Natural Gas Interstate and Intrastate Pipelines”. Accessed October 23, 2017. https://www.eia.gov/maps/layer_info-m.php.

Volk, Heather E., et al. 2013. “Traffic Related Air Pollution, Particulate Matter, and Autism”. *JAMA Psychiatry* 70 (1): 71–77.

Wilhoit, Larry R. 2012. “California Pesticide Use Data and Endangered Species”. Chap. Chapter 7 in *Pesticide Regulation and the Endangered Species Act*, ed. by Kenneth D. Racke et al., 1111:93–98. American Chemical Society.

Woolridge, Jeffrey. 2009. *Introductory Econometrics - A Modern Approach*. 4e. Mason, OH: South-Western Cengage Learning.

Woolridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second. Cambridge, Massachusetts: The MIT Press. ISBN: 9780262232197.

Xia, Y., et al. NCEP/EMC(2009). “NLDAS Primary Forcing Data L4 Hourly 0.125 x 0.125 degree V002”, ed. by David Mocko. Accessed April 30, 2018. NASA/GSFC/HSL, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).

Xu, Yiqing. 2017. “Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models”. *Political Analysis* 25:57–76.

A | Natural gas price elasticities and optimal cost recovery under consumer heterogeneity: Evidence from 300 million natural gas bills

A.1 Figures

Figure A.1: **California's 16 CEC climate zones** determine daily allowance within season

Notes: The shapefile underlying this map comes from the California Energy Commission (CEC). This map constitutes the CEC's climate-based building zones, which affect a number of energy policies, including households' baseline allowances. California Energy Commission 2015, 2017

Figure A.2: **Example bill:** PG&E residential natural gas bill

PG&E

ENERGY STATEMENT

www.pge.com/MyEnergy

Account No:	[redacted]
Statement Date:	12/25/2016
Due Date:	**01/17/2017**

Details of Gas Charges

11/24/2016 - 12/23/2016 (30 billing days)

Service For: [redacted]

Service Agreement ID: [redacted]

Rate Schedule: G1 T Residential Service

11/24/2016 – 11/30/2016	**Your Tier Usage**	**1**	**2**

Tier 1 Allowance	12.53 Therms	(7 days x 1.79 Therms/day)	
Tier 1 Usage	2.333330 Therms	@ $1.42692	$3.33
Gas PPP Surcharge ($0.10197 /Therm)			0.24
Berkeley Utility Users' Tax (7.500%)			0.25

12/01/2016 – 12/23/2016	**Your Tier Usage**	**1**	**2**

Tier 1 Allowance	41.17 Therms	(23 days x 1.79 Therms/day)	
Tier 1 Usage	7.666670 Therms	@ $1.36245	$10.45
Gas PPP Surcharge ($0.10197 /Therm)			0.78
Berkeley Utility Users' Tax (7.500%)			0.78

Total Gas Charges **$15.83**

Service Information

Meter #	[redacted]
Current Meter Reading	3,036
Prior Meter Reading	3,026
Difference	10
Multiplier	1.046971
Total Usage	10.000000 Therms
Baseline Territory	T
Serial	C

Gas Procurement Costs ($/Therm)

11/24/2016 - 11/30/2016	$0.45875
12/01/2016 - 12/23/2016	$0.39428

Notes: This 30-day bill for a PG&E customer (one of the authors) overlaps two calendar months in 2016: 7 days in November (24–30) and 23 days in December (01–23). Because PG&E's prices vary with the calendar month, PG&E needs to split consumption by calendar month. To achieve this task, PG&E assumes the customer consumed evenly across all days in the bill. Specifically, PG&E calculates that the customer consumed 10 therms and assigns the same amount of consumption to each day during the 30-day period. Thus, PG&E assigns $10 \times 7/30 \approx 2.33$ to November (the consumer spent 7 days in November in this 30-day bill) and $10 \times 23/30 \approx 7.67$ to December (the consumer spent 23 days in November in this 30-day bill).

Figure A.3: **PRISM:** Mean temperature raster for 15 June 2010

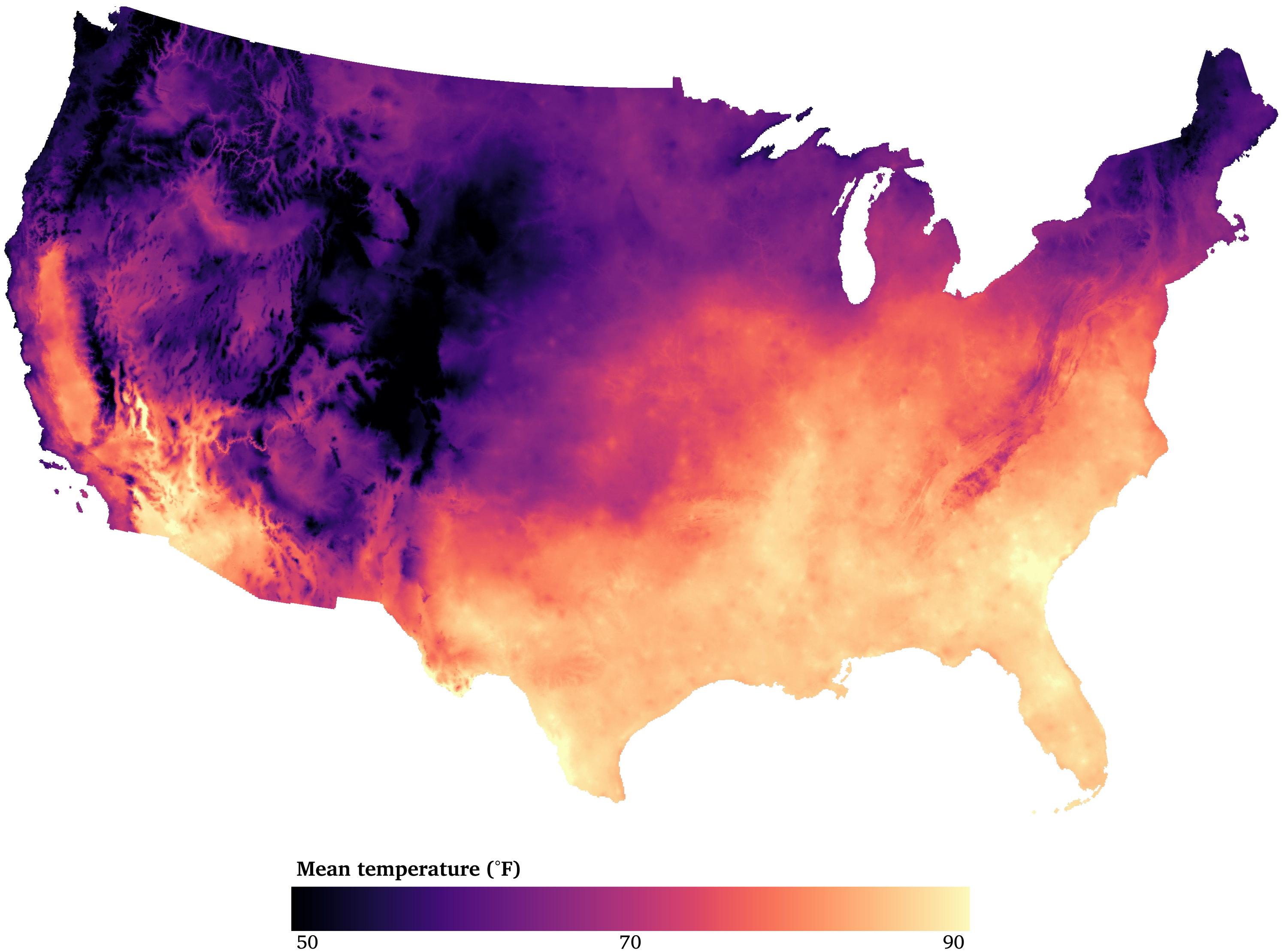
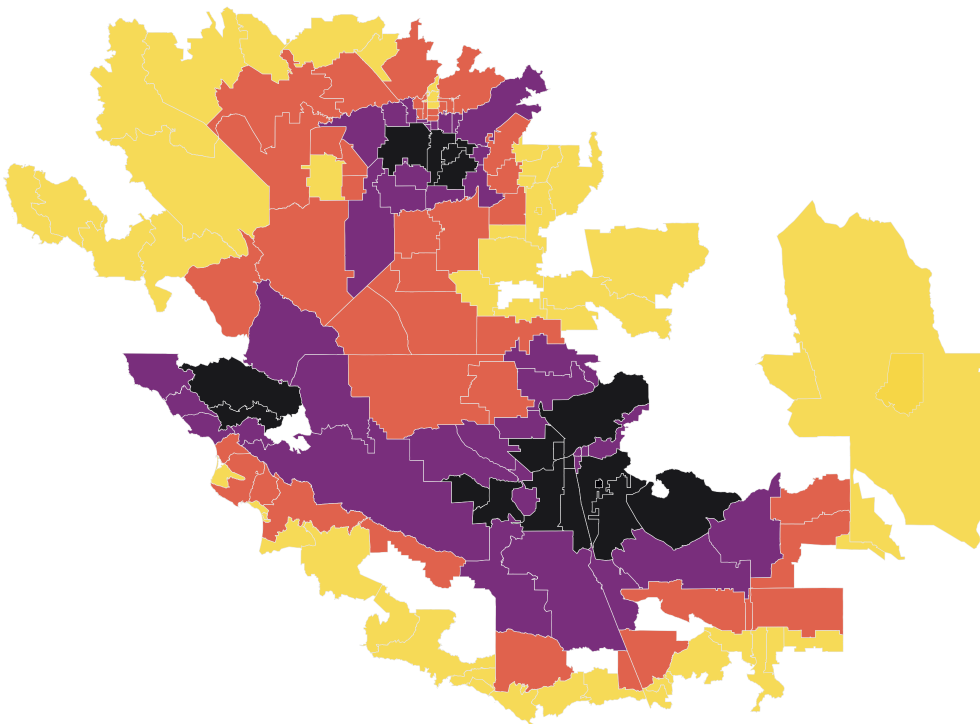


Figure A.4: **Expanding the study area:** Zip codes neighboring the study's zip codes



Zip-code group

Common Zips: zip codes served by both utilities

Neighbors 1: neighbors to Common Zips

Neighbors 2: neighbors to Neighbors 1

Neighbors 3: neighbors to Neighbors 2

Notes: This figure illustrates the four groups of zip codes referenced in Table A.16. The groups begin with *Common Zips*—the group in which each zip code receives natural gas service from both PG&E and SoCalGas—and expands by adding each group's immediately proximate neighbors. *E.g.*, *Neighbors 2* consists of all zip codes that neighbor a zip code in *Neighbors 1* (excluding those zip codes already included in another group).

A.2 Tables

Table A.1: **Price correlation**: Bivariate correlations between types of prices

	Type of Price				
	Marginal	Average	Avg. Mrg.	Baseline	Sim. mrg.
Marginal	1				
Average	0.8898	1			
Avg. Mrg.	0.8628	0.9421	1		
Baseline	0.7901	0.942	0.9202	1	
Sim. mrg.	0.8503	0.849	0.8174	0.781	1

Notes: *Avg.* or *average* price is the total bill divided by quantity. *Avg. Mrg.* or *average marginal* price denotes the quantity-weighted average of the household's marginal price. *Base* or *baseline* price refers to the price the household pays for its first unit (*therm*) of natural gas. *Sim. Mrg.* or *simulated marginal* price is the household's marginal price (using the relevant pricing regime) as a function of the household's historical consumption patterns (lagged bills 10 through 14).

Table A.2: **Testing the simulated instrument**:
Regressing marginal price on *simulated* marginal price

Dependent variable: Marginal price		
	(1)	(2)
Simulated marginal price	0.6425***	0.6444***
	(0.00435)	(0.00433)
Bill HDDs	T	T
Household FE	T	T
City month-of-sample FE	T	T
Lags used for sim. inst.	10–14	11–13
N	4,892,064	4,785,877

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Sim. Mrg.* or *simulated marginal* price is the household's marginal price (using the relevant pricing regime) as a function of the household's historical consumption patterns (lagged bills 10 through 14 or 11 through 13). As discussed in the empirical strategy section, the numbers of observations differ due to the lags required to calculate the *simulated instrument* for marginal price. *Significance levels:* *10%, **5%, ***1%.

Table A.3: **Comparing lags, second-stage results**: Marginal price with HH spot price IV

Dependent variable: Log(Consumption, daily avg.)

	Lag of Marginal Price				
	(1) 1 Lead	(2) No lag	(3) 1 Lag	(4) 2 Lags	(5) 3 Lags
Log(Price)	0.0480	−0.1121	−0.0223	−0.2098***	−0.1582**
instrumented	(0.0902)	(0.0762)	(0.0668)	(0.0706)	(0.0698)
First-stage F stat.	326.7	337.9	410.8	418.4	403.4
Bill HDDs	T	T	T	T	T
Household FE	T	T	T	T	T
City month-of-sample FE	T	T	T	T	T
N	5,501,467	5,754,088	5,754,088	5,754,085	5,754,079

Notes: With regard to lags: *No lag* refers to the price for the household's contemporaneous bill; *1 Lag* refers to the price on the household's previous bill; *etc. (HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Significance levels:* *10%, **5%, ***1%.

Table A.4: **Comparing lags, second-stage results**: Sim. marginal price with HH spot price IV

Dependent variable: Log(Consumption, daily avg.)

	Lag of Simulated Marginal Price				
	(1) 1 Lead	(2) No lag	(3) 1 Lag	(4) 2 Lags	(5) 3 Lags
Log(Price)	0.0317	−0.0549	0.0329	−0.1705**	−0.1596**
instrumented	(0.0899)	(0.0718)	(0.0626)	(0.0698)	(0.0720)
First-stage F stat.	354.7	379.6	393.2	369.9	332.1
Bill HDDs	T	T	T	T	T
Household FE	T	T	T	T	T
City month-of-sample FE	T	T	T	T	T
N	4,778,382	4,892,064	4,785,877	4,682,526	4,590,790

Notes: With regard to lags: *No lag* refers to the price for the household's contemporaneous bill; *1 Lag* refers to the price on the household's previous bill; *etc. Sim. Mrg.* or *simulated marginal* price is the household's marginal price (using the relevant pricing regime) as a function of the household's historical consumption patterns (lagged bills 10 through 14). *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Significance levels:* *10%, **5%, ***1%.

Table A.5: **Comparing lags, second-stage results**: Avg. marginal price with HH spot price IV

Dependent variable: Log(Consumption, daily avg.)

	Lag of Average Marginal Price				
	(1) 1 Lead	(2) No lag	(3) 1 Lag	(4) 2 Lags	(5) 3 Lags
Log(Price)	0.0432	−0.0853	−0.0313	−0.1734***	−0.1356**
instrumented	(0.0745)	(0.0618)	(0.0568)	(0.0585)	(0.0585)
First-stage F stat.	969.4	1,036.4	1,275.3	1,311.0	1,306.1
Bill HDDs	T	T	T	T	T
Household FE	T	T	T	T	T
City month-of-sample FE	T	T	T	T	T
N	5,501,467	5,754,088	5,754,088	5,754,085	5,754,079

Notes: With regard to lags: *No lag* refers to the price for the household's contemporaneous bill; *1 Lag* refers to the price on the household's previous bill; *etc. Sim. Mrg.* or *simulated marginal* price is the household's marginal price (using the relevant pricing regime) as a function of the household's historical consumption patterns (lagged bills 10 through 14). *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Significance levels:* *10%, **5%, ***1%.

Table A.6: **Comparing lags, second-stage results**: Avgerage price with HH spot price IV

Dependent variable: Log(Consumption, daily avg.)

	Lag of Average Price				
	(1) 1 Lead	(2) No lag	(3) 1 Lag	(4) 2 Lags	(5) 3 Lags
Log(Price)	0.0515	−0.1244	−0.0177	−0.2312***	−0.1680**
instrumented	(0.0972)	(0.0805)	(0.0730)	(0.0760)	(0.0749)
First-stage F stat.	679.1	725.8	884.4	899.4	923.7
Bill HDDs	T	T	T	T	T
Household FE	T	T	T	T	T
City month-of-sample FE	T	T	T	T	T
N	5,501,467	5,754,088	5,754,088	5,754,085	5,754,079

Notes: With regard to lags: *No lag* refers to the price for the household's contemporaneous bill; *1 Lag* refers to the price on the household's previous bill; *etc. Avg.* or *average* price is the total bill divided by quantity. *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Significance levels:* *10%, **5%, ***1%.

Table A.7: **Comparing lags, second-stage results**: Baseline price with HH spot price IV

Dependent variable: Log(Consumption, daily avg.)					
	Lag of Baseline Price				
	(1) 1 Lead	(2) No lag	(3) 1 Lag	(4) 2 Lags	(5) 3 Lags
Log(Price)	0.0420	−0.1164*	−0.0066	−0.2030***	−0.1396**
instrumented	(0.0839)	(0.0684)	(0.0637)	(0.0650)	(0.0630)
First-stage F stat.	1,085.3	1,143.4	1,241.8	1,333.2	1,533.2
Bill HDDs	T	T	T	T	T
Household FE	T	T	T	T	T
City month-of-sample FE	T	T	T	T	T
N	5,501,467	5,754,088	5,754,088	5,754,085	5,754,079

Notes: With regard to lags: *No lag* refers to the price for the household's contemporaneous bill; *1 Lag* refers to the price on the household's previous bill; *etc. Base* or *baseline* price refers to the price the household pays for its first unit (*therm*) of natural gas. *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Significance levels:* *10%, **5%, ***1%.

Table A.8: **Heterogeneity by season or income**:
Second-stage results, instrumenting average price with HH spot price

Dependent variable: Log(Consumption, daily avg.)

	Average Price			
	Split by **Season**		Split by **CARE (Income)**	
	(1) Summer	(2) Winter	(3) CARE	(4) Non-CARE
Log(Price)	0.0579*	−0.4694***	−0.2650***	−0.1557**
instrumented	(0.0316)	(0.1586)	(0.0834)	(0.0740)
First-stage F stat.	765.7	223.4	814.7	745.8
Bill HDDs	T	T	T	T
Household FE	T	T	T	T
City month-of-sample FE	T	T	T	T
N	3,065,917	2,688,168	2,435,135	3,318,950

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. *Summer* includes April through September. *Winter* includes October through March. *CARE* households participate in the California Alternative Rates for Energy (CARE) program. CARE targets low-income households and provides a 20 percent discount on natural gas bills. We estimate the heterogeneity results by splitting the sample along the dimension(s) of heterogeneity and then individually estimating the models. *Avg.* or *average* price is the total bill divided by quantity. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Significance levels:* *10%, **5%, ***1%.

Table A.9: **Heterogeneity by season and income**:
Second-stage results, instrumenting average price with HH spot price

Dependent variable: Log(Consumption, daily avg.)

	Average Price			
	(1) Summer CARE	(2) Summer Non-CARE	(3) Winter CARE	(4) Winter Non-CARE
Log(Price)	0.0495	0.0828**	−0.6106***	−0.3971**
instrumented	(0.0384)	(0.0359)	(0.1570)	(0.1687)
First-stage F stat.	691.5	591.9	212.7	184.8
Bill HDDs	T	T	T	T
Household FE	T	T	T	T
City month-of-sample FE	T	T	T	T
N	1,293,144	1,772,773	1,141,991	1,546,177

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Summer* includes April through September. *Winter* includes October through March. *CARE* households participate in the California Alternative Rates for Energy (CARE) program. CARE targets low-income households and provides a 20 percent discount on natural gas bills. We estimate the heterogeneity results by splitting the sample along the dimension(s) of heterogeneity and then individually estimating the models. *Avg.* or *average* price is the total bill divided by quantity. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Significance levels:* *10%, **5%, ***1%.

Table A.10: **First-stage results**:
Robustness to specification: Marginal price instrumented with spot price

Dependent variable: Log(Marginal price)

	(1)	(2)	(3)	(4)
Spot price	0.3398***	0.3679***	0.3806***	0.3955***
	(0.0757)	(0.0774)	(0.0798)	(0.0547)
Spot price × SoCalGas	0.7858***	0.7868***	0.7856***	0.7385***
	(0.0300)	(0.0299)	(0.0302)	(0.0378)
First-stage F stat.	416.1	418.4	415.2	367.0
Bill HDDs	F	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	F	F
City by week-of-sample FE	F	F	T	F
Zip by week-of-sample FE	F	F	F	T
N	5,754,085	5,754,085	5,754,085	5,754,085

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *(HH) Spot price* refers to the weekly average spot price for natural gas at Louisiana's Henry Hub in the week preceding the utility's price change. *Significance levels:* *10%, **5%, ***1%.

Table A.11: **Second-stage results**:
Robustness to specification: Marginal price instrumented with spot price

Dependent variable: Log(Consumption, daily avg.)

	(1)	(2)	(3)	(4)
Log(Marginal price) *instrumented*	-0.3623^{***} (0.0854)	-0.2098^{***} (0.0706)	-0.1705^{***} (0.0621)	-0.1495^{**} (0.063)
First-stage F stat.	416.1	418.4	415.2	367.0
Bill HDDs	F	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	F	F
City by week-of-sample FE	F	F	T	F
Zip by week-of-sample FE	F	F	F	T
N	5,754,085	5,754,085	5,754,085	5,754,085

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Significance levels:* *10%, **5%, ***1%.

Table A.12: **Second-stage results**:
Robustness to specification: Average price instrumented with spot price

Dependent variable: Log(Consumption, daily avg.)

	(1)	(2)	(3)	(4)
Log(Average price) *instrumented*	-0.4076^{***} (0.0911)	-0.2312^{***} (0.076)	-0.1891^{***} (0.067)	-0.1574^{**} (0.0656)
First-stage F stat.	897.5	899.4	881.1	661.1
Bill HDDs	F	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	F	F
City by week-of-sample FE	F	F	T	F
Zip by week-of-sample FE	F	F	F	T
N	5,754,085	5,754,085	5,754,085	5,754,085

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Avg.* or *average* price is the total bill divided by quantity. *Significance levels:* *10%, **5%, ***1%.

Table A.13: **Second-stage results**:
Robustness to specification: Avg. mrg. price instrumented with spot price

Dependent variable: Log(Consumption, daily avg.)

	(1)	(2)	(3)	(4)
Log(Avg. marginal price)	-0.2951^{***}	-0.1734^{***}	-0.1529^{***}	-0.1330^{**}
instrumented	(0.0697)	(0.0585)	(0.0514)	(0.0549)
First-stage F stat.	1,299.9	1,311.0	1,275.8	780.6
Bill HDDs	F	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	F	F
City by week-of-sample FE	F	F	T	F
Zip by week-of-sample FE	F	F	F	T
N	5,754,085	5,754,085	5,754,085	5,754,085

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Avg. Mrg.* or *average marginal* price denotes the quantity-weighted average of the household's marginal price. *Significance levels:* *10%, **5%, ***1%.

Table A.14: **Second-stage results**:
Robustness to specification: Baseline price instrumented with spot price

Dependent variable: Log(Consumption, daily avg.)

	(1)	(2)	(3)	(4)
Log(Simulated mrg. price)	-0.3148^{***}	-0.1705^{**}	-0.1310^{**}	-0.1025
instrumented	(0.0843)	(0.0698)	(0.0602)	(0.0675)
First-stage F stat.	368.9	369.9	331.3	181.9
Bill HDDs	F	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	F	F
City by week-of-sample FE	F	F	T	F
Zip by week-of-sample FE	F	F	F	T
N	4,682,526	4,682,526	4,682,526	4,682,526

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Sim. Mrg.* or *simulated marginal* price is the household's marginal price (using the relevant pricing regime) as a function of the household's historical consumption patterns (lagged bills 10 through 14). *Significance levels:* *10%, **5%, ***1%.

Table A.15: **Second-stage results**:
Robustness to specification: Baseline price instrumented with spot price

Dependent variable: Log(Consumption, daily avg.)

	(1)	(2)	(3)	(4)
Log(Baseline price)	-0.3643^{***}	-0.2030^{***}	-0.1653^{***}	-0.1376^{**}
instrumented	(0.077)	(0.065)	(0.0576)	(0.0572)
First-stage F stat.	1,322.9	1,333.2	1,187.3	762.5
Bill HDDs	F	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	F	F
City by week-of-sample FE	F	F	T	F
Zip by week-of-sample FE	F	F	F	T
N	5,754,085	5,754,085	5,754,085	5,754,085

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Base* or *baseline* price refers to the price the household pays for its first unit (*therm*) of natural gas. *Significance levels:* *10%, **5%, ***1%.

Table A.16: **Second-stage results**: Extending the set of zip codes to neighboring zip codes

Dependent variable: Log(Consumption, daily avg.)				
	Marginal Price			
	(1) Common Zips	(2) Neighbors 1	(3) Neighbors 2	(4) Neighbors 3
Log(Marginal price)	-0.2098^{***}	-0.1896^{***}	-0.1241^{***}	-0.0946^{***}
instrumented	(0.0706)	(0.049)	(0.0401)	(0.0357)
First-stage F stat.	418.4	713.0	735.8	1,182.9
Bill HDDs	T	T	T	T
Household FE	T	T	T	T
City by month-of-sample FE	T	T	T	T
Levels of neighboring zip codes	0	1	2	3
N	5,754,085	11,679,371	19,629,128	28,277,567

Notes: *Common zips* refers the set of zip codes in which each zip code receives natural gas from both PG&E and SoCalGas. *Neighbors 1* includes the *common zips* and the zip codes that immediately neighbor the common zips. *Neighbors 2* adds the neighbors of these neighbors (adding the neighbors of *Neighbors 1*). *Neighbors 3* adds the neighbors of *Neighbors 2*. Figure A.4 depicts these sets of zip codes. Each column denotes a separate regression. Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle (the level at which price varies). All regressions include heating degree days (HDDs) within the households' billing period. Each price in the table is the second lag of price, *i.e.*, the prices from two bills prior to the current bill. *Significance levels:* *10%, **5%, ***1%.

Table A.17: **Billing data description:** Columns within the billing data

Feature name	**Description**
Account ID	Unique identifier for household account with the utility
Premise ID	Unique physical-location based identifier
Prior read date	Effectively the start date of the bill
Current read date	Effectively the end date of the bill
Gas rate schedule	Classifies type of customer (and the customer's price regime)
Gas usage	Volume of gas consumed during billing period (in therms)
Bill revenue	Total bill charged to household for the current billing period
Climate band	California Public Utility Commission-based climate region
Service address 9-digit zip	Household's 9-digit zip code
Service start date	Date on which the household began service
Service stop date	Date on which the household ended service

A.3 Data appendix

A.3.1 Calculating bills

As discussed in the body of the paper, the majority of bills do not line up with calendar months. Consequently, households' billing periods do not line up with utilities' monthly changes in price (or with changes in daily allowances resulting from changes in seasons). Thus, a single bill will typically span multiple price regimes. The two utilities deal with change in price in subtly different ways. This "problem" results from the fact that neither utility knows households' *daily* consumption.

PG&E When a PG&E customer's bill spans multiple calendar months (price regimes), PG&E calculates individual bills for each month. However, because PG&E does not know the daily consumption levels, they assume constant daily consumption throughout the billing period.

SoCalGas In the case that a SoCalGas customer's bill spans multiple calendar months (price regimes), SoCalGas computes time-weighted average prices (and allowances) by aggregating the prices and allowances from the calendar months by the number of days the bill spent in each month.

A.3.2 Data work

In this section, we describe the exclusion and cleaning choices that we made while preparing the data for analysis. Our R scripts are available upon request, though the data themselves cannot be shared due to agreements with the utilities and the IRB.
Exclusions:

- We omitted SoCalGas price data from advice letters 3644, 3680, 3695, 3807, 4053, and 4061, as they were updated by letters 3660, 3697, 3697, 3810, 4055, and 4069, respectively.
- We dropped pre-2008 data (PG&E and prices/allowances), as SoCalGas did not share billing data for pre-2009 bills.
- We trimmed the shortest 2.5% and longest 2.5% bills (resulted in keeping bills of length between 28-34 days). We did this to omit the first or last bills for a household and bills that were irregular for any other reasons. We applied this requirement of 28–34 days to the current bill and the first through the third lagged bills, because we consider the effect of lagged prices on contemporaneous consumption.
- We dropped bills missing any critical information: number of therms (quantity), revenue, *etc.*
- We dropped bills outside the central 99% of data (*i.e.*, the bill's revenue or volume fall in the bottom 0.5% or in the top 0.5%). Our main results apply this rule for the contemporaneous and the first three lagged bills.

- We dropped bills whose total revenue we could not predict within five percent (using known prices, quantities, and discounts).
- We dropped bills for exactly zero therms.

CARE status While the datasets presumably denoted CARE (California Alternate Rates for Energy) households, we found many households not denoted as CARE households whose charges were consistent with CARE pricing (*i.e.*, charges were 80 percent of the standard tariffs). We classified these households as CARE households.

B | Are our hopes too high? Testing cannabis legalization's potential to reduce criminalization

B.1 Tables

Table B.1: **Effect of cannabis legalization on drug-related offenses**: Difference-in-differences results

	Dependent variable			
	(1) # of incidents	(2) Log(# of incidents)	(3) # of cannabis incidents	(4) # of non-cannabis incidents
Legalization *indicator*	40.5858*** (6.83)	0.8979*** (0.143)	11.9676*** (3.71)	15.3149*** (2.46)
City FE	T	T	T	T
Year-Week FE	T	T	T	T
N	11,723	11,723	11,723	11,723

Notes: Each column denotes a separate regression with a different dependent variable. The observational unit is city i in week of sample t. Errors are clustered within a city in a year. The mean number of drug-related police incidents before Colorado's legalization was 41.62 (20.26 cannabis-related offenses and 16.03 non-cannabis related offenses). *Significance levels:* *10%, **5%, ***1%.

B.2 Model appendix

B.2.1 Consumer comparative statics

Drawing upon the consumer's first-order conditions in (2.6) and (2.7) (call this system of equations F), the Hessian matrix of second derivatives is

$$\mathscr{H} = D_{\boldsymbol{x}}F = \begin{bmatrix} U_{11}(x_1, x_2) & U_{12}(x_1, x_2) \\ U_{21}(x_1, x_2) & U_{22}(x_1, x_2) \end{bmatrix} \tag{B.1}$$

Assuming that consumers indeed maximize their utility on well-behaved utility functions, the determinant of $\mathscr{H}$ is positive.

By taking partial derivatives of (2.6) and (2.7) with respect to the five exogenous[1] parameters of interest ($\boldsymbol{p} = \{e_1, e_2, \Gamma_1, \Gamma_2, \kappa\}$), we get the matrix

$$D_{\boldsymbol{p}}F = \begin{bmatrix} -\kappa\Gamma_1 n_1' & 0 & -\kappa n_1 & 0 & -n_1\Gamma_1 \\ 0 & -\kappa\Gamma_2 n_2' & 0 & -\kappa n_2 & -n_2\Gamma_2 \end{bmatrix} \tag{B.2}$$

Assuming $\mathscr{H}$ is non-zero in the vicinity of the equilibrium—which follows from assuming the consumer's problem leads to a maximum—the implicit function theorem implies

$$\begin{bmatrix} \frac{\partial x_1}{\partial e_1} & \frac{\partial x_1}{\partial e_2} & \frac{\partial x_1}{\partial \Gamma_1} & \frac{\partial x_1}{\partial \Gamma_1} & \frac{\partial x_1}{\partial \kappa} \\ \frac{\partial x_2}{\partial e_1} & \frac{\partial x_2}{\partial e_2} & \frac{\partial x_2}{\partial \Gamma_1} & \frac{\partial x_2}{\partial \Gamma_1} & \frac{\partial x_2}{\partial \kappa} \end{bmatrix} = -\mathscr{H}^{-1} D_{\boldsymbol{p}}F \tag{B.3}$$

Now define the inverse of the determinant of $\mathscr{H}$ as $\mathscr{A}$, *i.e.*,

$$\mathscr{H}^{-1} = \frac{1}{\det \mathscr{H}} \begin{bmatrix} U_{22} & -U_{12} \\ -U_{12} & U_{22} \end{bmatrix} = \mathscr{A} \begin{bmatrix} U_{22} & -U_{12} \\ -U_{12} & U_{22} \end{bmatrix} \tag{B.4}$$

then (B.3) implies

$$\begin{gathered} \begin{bmatrix} \frac{\partial x_1}{\partial e_1} & \frac{\partial x_1}{\partial e_2} & \frac{\partial x_1}{\partial \Gamma_1} & \frac{\partial x_1}{\partial \Gamma_1} & \frac{\partial x_1}{\partial \kappa} \\ \frac{\partial x_2}{\partial e_1} & \frac{\partial x_2}{\partial e_2} & \frac{\partial x_2}{\partial \Gamma_1} & \frac{\partial x_2}{\partial \Gamma_1} & \frac{\partial x_2}{\partial \kappa} \end{bmatrix} = \\ \mathscr{A} \begin{bmatrix} U_{22}\kappa\Gamma_1 n_1' & -U_{12}\kappa\Gamma_2 n_2' & U_{22}\kappa n_1 & -U_{12}\kappa n_2 & U_{22}n_1\Gamma_1 - U_{12}n_2\Gamma_2 \\ -U_{21}\kappa\Gamma_1 n_1' & U_{11}\kappa\Gamma_2 n_2' & -U_{21}\kappa n_1 & U_{11}\kappa n_2 & U_{11}n_2\Gamma_2 - U_{21}n_1\Gamma_1 \end{bmatrix} \end{gathered} \tag{B.5}$$

[1] *Exogeneous* from the consumer's perspective.

B.2.2 Officer optimization

Substituting $\partial x_i/\partial e_i$ and $\partial x_i/\partial e_j$ from (2.8) and (2.9) into the officer's first-order conditions in (2.14–2.16) yields

$$\mathscr{L}_{e_1} = \gamma_1 \,(n_1'\, x_1 + n_1\, \mathscr{A}\, U_{22}\, \kappa\, \Gamma_1\, n_1') + \gamma_2\, n_2\, \mathscr{A}\, U_{21}\, \kappa\, \Gamma_1\, n_1' - \lambda = 0 \tag{B.6}$$

$$\mathscr{L}_{e_2} = \gamma_1\, n_1\, \mathscr{A}\, U_{12}\, \kappa\, \Gamma_2\, n_2' + \gamma_2 \,(n_2'\, x_2 + n_2\, \mathscr{A}\, U_{11}\, \kappa\, \Gamma_2\, n_2') - \lambda = 0 \tag{B.7}$$

$$\mathscr{L}_{\lambda} = E - e_1 - e_2 = 0 \tag{B.8}$$

In addition, the first-order conditions from the officer's constrained maximization problem generate the bordered Hessian

$$\mathscr{H}_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & \mathscr{L}_{11} & \mathscr{L}_{12} \\ 1 & \mathscr{L}_{21} & \mathscr{L}_{22} \end{bmatrix} \tag{B.9}$$

which, by the second-order condition of the officer's constrained maximization, implies $\det \mathscr{H}_{\mathscr{B}} = \mathscr{L}_{12} + \mathscr{L}_{21} - \mathscr{L}_{11} - \mathscr{L}_{22} > 0$.

Defining the system of equations given in (B.6–B.8) as G, we can again apply the implicit function. I will assume that the second derivatives of the consumer's utility function are constant in the area surrounding the equilibrium. First, take derivatives of G with respect to the three endogenous variables $\boldsymbol{e} = \{e_1, e_2, \lambda\}$, *i.e.*, $D_{\boldsymbol{e}}G$:

$$\frac{\partial \mathscr{L}_{e_1}}{\partial e_1} = \gamma_1 \left\{ n_1'' x_1 + n_1' \frac{\partial x_1}{\partial e_1} + \mathscr{A} U_{22} \kappa \Gamma_1 \left[(n_1')^2 + n_1 n_1'' \right] \right\} - \gamma_2 n_2 \mathscr{A} U_{21} \kappa \Gamma_1 n_1''$$

$$\frac{\partial \mathscr{L}_{e_1}}{\partial e_2} = \gamma_1 n_1' \frac{\partial x_1}{\partial e_1} - \gamma_2 n_2' \mathscr{A} U_{21} \kappa \Gamma_1 n_1'$$

$$\frac{\partial \mathscr{L}_{e_1}}{\partial \lambda} = -1$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial e_1} = -\gamma_1 n_1' \mathscr{A} U_{12} \kappa \Gamma_2 n_2' + \gamma_2 n_2' \frac{\partial x_2}{\partial e_1}$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial e_2} = -\gamma_1 n_1 \mathscr{A} U_{12} \kappa \Gamma_2 n_2'' + \gamma_2 \left\{ n_2'' x_2 + n_2' \frac{\partial x_2}{\partial e_2} + \mathscr{A} U_{11} \kappa \Gamma_2 \left[(n_2')^2 + n_2 n_2'' \right] \right\}$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial \lambda} = -1$$

$$\frac{\partial \mathscr{L}_{\lambda}}{\partial \lambda} = -1$$

$$\frac{\partial \mathscr{L}_\lambda}{\partial \lambda} = -1$$

$$\frac{\partial \mathscr{L}_\lambda}{\partial \lambda} = 0$$

Defining now define the exogenous parameters $\boldsymbol{m} = \{\gamma_1, \gamma_2, \Gamma_1, \Gamma_2, \kappa\}$ and differentiate G with respect to $\boldsymbol{m}$, *i.e.*, $D_{\boldsymbol{m}}G$:

$$\frac{\partial \mathscr{L}_{e_1}}{\partial \gamma_1} = n_1' \left(x_1 + n_1 \mathscr{A} U_{22} \kappa \Gamma_1\right)$$

$$\frac{\partial \mathscr{L}_{e_1}}{\partial \gamma_2} = -n_2 \mathscr{A} U_{21} \kappa \Gamma_1 n_1'$$

$$\frac{\partial \mathscr{L}_{e_1}}{\partial \Gamma_1} = \mathscr{A} \kappa n_1' \left(\gamma_1 n_1 U_{22} - \gamma_2 n_2 U_{21}\right)$$

$$\frac{\partial \mathscr{L}_{e_1}}{\partial \Gamma_2} = 0$$

$$\frac{\partial \mathscr{L}_{e_1}}{\partial \kappa} = \gamma_1 n_1 \mathscr{A} U_{22} \Gamma_1 n_1' - \gamma_2 n_2 \mathscr{A} U_{21} \Gamma_1 n_1'$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial \gamma_1} = -n_1 \mathscr{A} U_{12} \kappa \Gamma_2 n_2'$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial \gamma_2} = n_2' \left(x_2 + \mathscr{A} U_{11} \kappa \Gamma_2\right)$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial \Gamma_1} = 0$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial \Gamma_2} = \mathscr{A} \kappa n_2' (\gamma_2 n_2 U_{11} - \gamma_1 n_1 U_{12})$$

$$\frac{\partial \mathscr{L}_{e_2}}{\partial \kappa} = -\gamma_1 n_1 \mathscr{A} U_{12} \Gamma_2 n_2' + \gamma_2 n_2 \mathscr{A} U_{11} \Gamma_2 n_2'$$

$$\frac{\partial \mathscr{L}_\lambda}{\partial m} = 0, \forall m \in \boldsymbol{m}$$

Then, by the implicit function theorem,

$$\begin{bmatrix} \frac{\partial e_1}{\partial \gamma_1} & \frac{\partial e_1}{\partial \gamma_2} & \frac{\partial e_1}{\partial \Gamma_1} & \frac{\partial e_1}{\partial \Gamma_1} & \frac{\partial e_1}{\partial \kappa} \\ \frac{\partial e_2}{\partial \gamma_1} & \frac{\partial e_2}{\partial \gamma_2} & \frac{\partial e_2}{\partial \Gamma_1} & \frac{\partial e_2}{\partial \Gamma_1} & \frac{\partial e_2}{\partial \kappa} \end{bmatrix} = -[D_e G]^{-1} D_m G \tag{B.10}$$

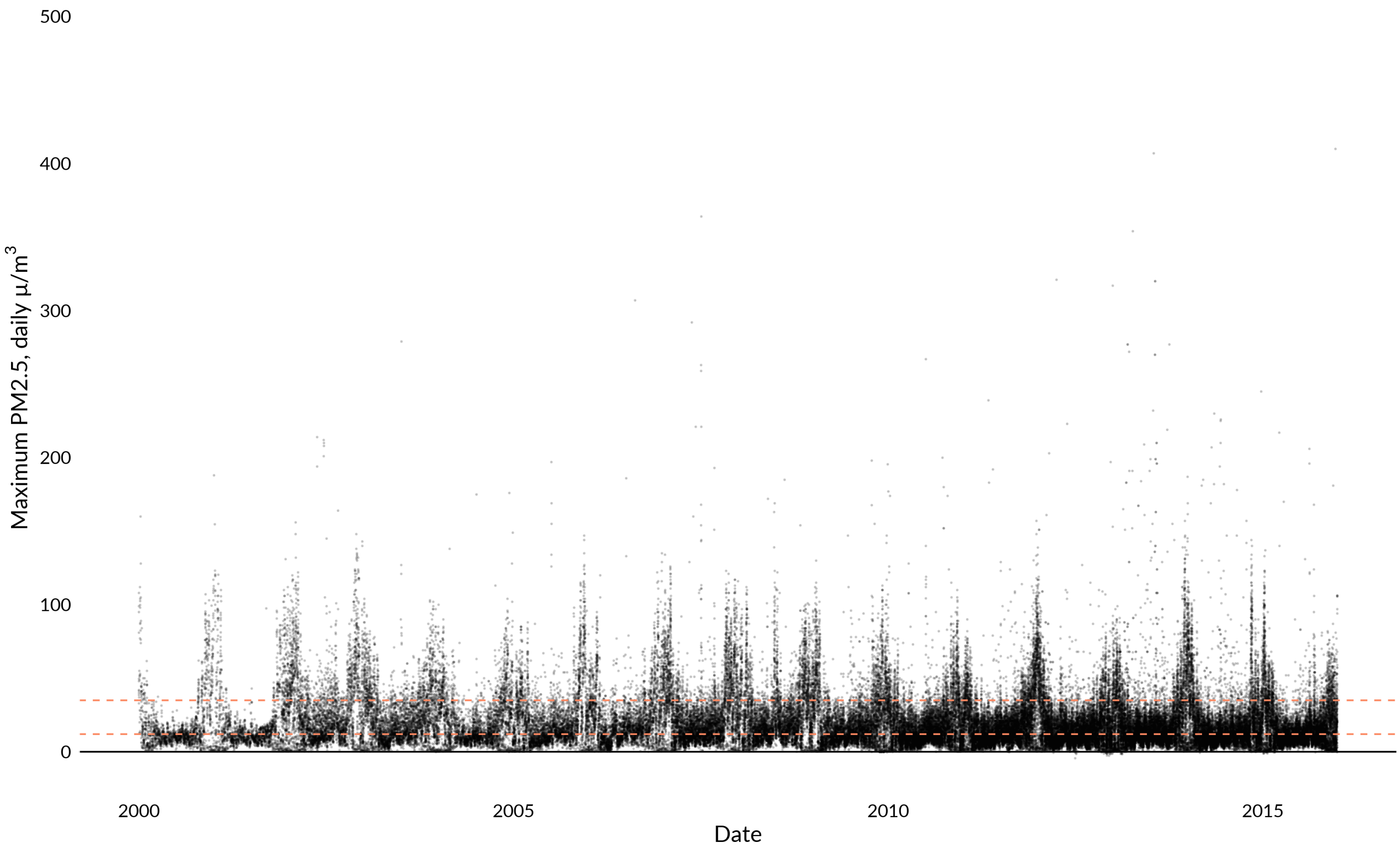
$$= -\mathscr{B} \begin{bmatrix} -1 & 1 & \mathscr{L}_{22} - \mathscr{L}_{12} \\ 1 & -1 & \mathscr{L}_{11} - \mathscr{L}_{21} \\ \mathscr{L}_{22} - \mathscr{L}_{21} & \mathscr{L}_{11} - \mathscr{L}_{12} & \mathscr{L}_{11}\mathscr{L}_{22} - \mathscr{L}_{12}\mathscr{L}_{21} \end{bmatrix} D_m G \tag{B.11}$$

where $\mathscr{B} = \det D_e G$, which is positive by the second-order conditions of constrained maximization, and $\mathscr{L}_{ij}$ denotes $\partial \mathscr{L}_{e_i}/\partial e_j$.

C | Do aerially applied pesticides affect local air quality? Empirical evidence from California's San Joaquin Valley

C.1 Figures

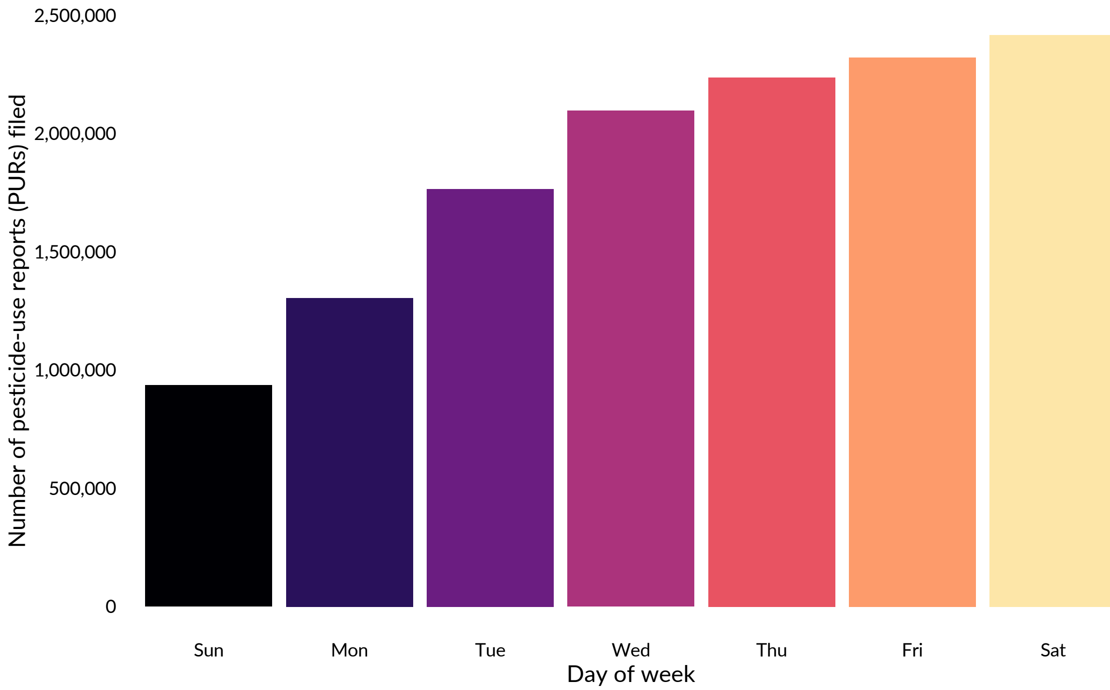
Figure C.1: **EPA PM2.5 records:** Daily maximum PM2.5 concentration at each monitor, 2000–2015



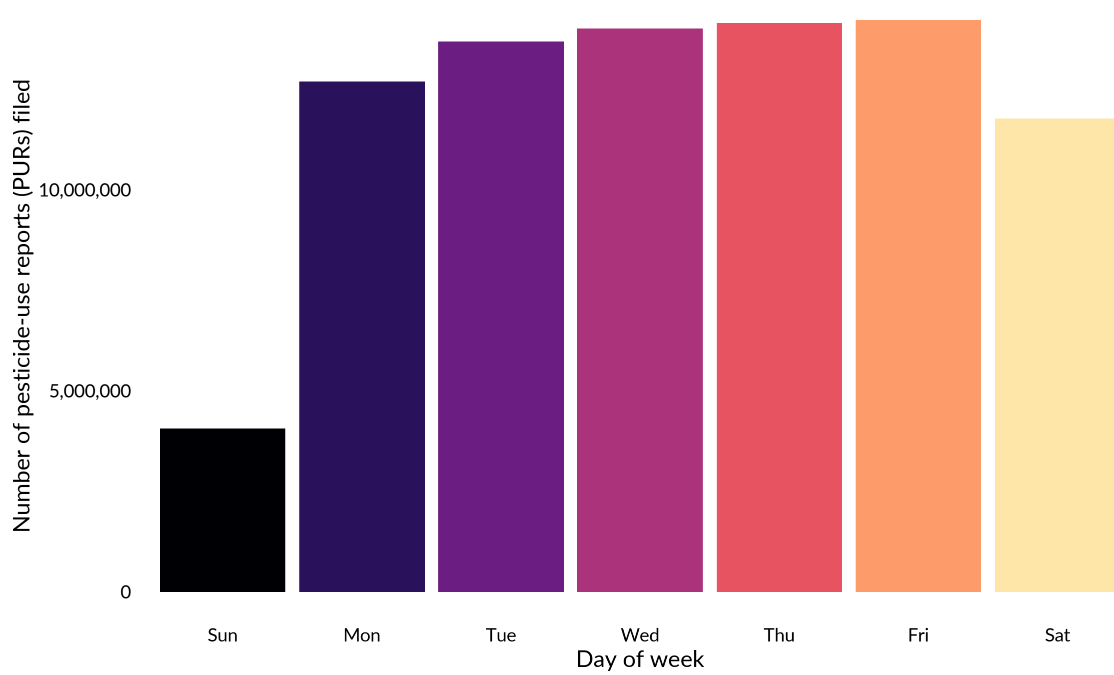
Notes: Each point in this figure represents the maximum PM2.5 on the given day (*x* axis) for a specific monitor. The two dashed horizontal lines denote two different primary National Ambient Air Quality Standards (NAAQS), established January 15, 2013. The lower line establishes the standard (12.0 $\mu g/m^3$) for the the 3-year arithmetic mean. The higher line marks the standard (35 $\mu g/m^3$) for the 3-year mean of the 98th percentile (U.S. E.P.A. 2013).

Figure C.2: **Number of pesticide applications:** 2000–2015, by day of week

(a) **Aerially applied pesticides**



(b) **Ground-applied pesticides**



Notes: These figures depict the number PURs filed by day of week in the five study counties between 2000 and 2015, separating the pesticide applications by aerial applications (top) and ground applications (bottom).*Source:* Author using data from California D.P.R. 2013.

C.2 Tables

Table C.1: **Increases in mean PM2.5**: Same-day, aerially applied pesticides, Windsorized

Dependent variable: Mean daily PM2.5 level

	(1)	(2)	(3)	(4)	(5)
Tons of pesticide (**a**) within 1.5km	0.2575***	0.1455**	0.2066*	0.2139**	0.2545**
	(0.0915)	(0.0588)	(0.1094)	(0.0984)	(0.1180)
Tons of pesticide (**b**) between 1.5km and 3km	0.0613	0.0295	0.0588	0.0482	0.0660
	(0.1421)	(0.0787)	(0.0868)	(0.1047)	(0.1160)
Tons of pesticide (**c**) between 3km and 25km	0.0413	0.1047	0.0625	0.0413	0.0124
	(0.0788)	(0.0664)	(0.0645)	(0.0790)	(0.0767)
Monitor FE	T	T	T	T	T
Day-of-sample FE	T	F	F	F	F
Week-of-sample FE	F	T	F	F	F
Month-of-sample FE	F	F	T	F	F
Week-of-year FE	F	F	F	T	F
Month-of-year FE	F	F	F	F	T
Day-of-week FE	F	F	F	T	T
Year FE	F	F	F	T	T
N	26,242	26,242	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Windsorization* replaces any value *above* the 97.5^{th} percentile of a variable with the 97.5^{th} percentile. Windsorization attempts to limit the influence of extreme values within the independent variables(s). I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.*, '(**a**)') reference labeled areas in the figures associated with these results.

Table C.2: **Increases in max. PM2.5**: Same-day, aerially applied pesticides, Windsorized

Dependent variable: Maximum daily PM2.5 level

	(1)	(2)	(3)	(4)	(5)
Tons of pesticide (**a**)	0.3400***	0.1466***	0.2006**	0.1730**	0.2140**
within 1.5km	(0.1026)	(0.0484)	(0.0975)	(0.0841)	(0.0994)
Tons of pesticide (**b**)	0.1140	0.0949	0.1327	0.1188	0.1384
between 1.5km and 3km	(0.1792)	(0.0968)	(0.0997)	(0.1223)	(0.1346)
Tons of pesticide (**c**)	0.0218	0.0845	0.0349	0.0030	−0.0321
between 3km and 25km	(0.0862)	(0.0740)	(0.0752)	(0.0804)	(0.0802)
Monitor FE	T	T	T	T	T
Day-of-sample FE	T	F	F	F	F
Week-of-sample FE	F	T	F	F	F
Month-of-sample FE	F	F	T	F	F
Week-of-year FE	F	F	F	T	F
Month-of-year FE	F	F	F	F	T
Day-of-week FE	F	F	F	T	T
Year FE	F	F	F	T	T
N	26,242	26,242	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Windsorization* replaces any value *above* the 97.5th percentile of a variable with the 97.5th percentile. Windsorization attempts to limit the influence of extreme values within the independent variables(s). I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.*, '(**a**)') reference labeled areas in the figures associated with these results.

Table C.3: **Increases in mean PM2.5 from aerial pesticides**: Wind-variation results, Windsorized

Dependent variable: Mean daily PM2.5 level	(1)	(2)	(3)
Tons of pesticide (**a**) within 1.5km; Upwind	0.2954**	0.2557***	0.2607**
	(0.1504)	(0.0856)	(0.1094)
Tons of pesticide (**b**) within 1.5km; Orthogonal	−0.0080	−0.4481***	−0.0936
	(0.1395)	(0.0482)	(0.1039)
Tons of pesticide (**c**) within 1.5km; Downwind	0.3131**	0.2040**	0.2123
	(0.1549)	(0.0835)	(0.2183)
Tons of pesticide (**d**) between 1.5km and 3km; Upwind	0.0772	0.0704	0.1267
	(0.1327)	(0.1029)	(0.1018)
Tons of pesticide (**e**) between 1.5km and 3km; Orthogonal	0.0582	−0.0267	−0.0055
	(0.2065)	(0.1400)	(0.1100)
Tons of pesticide (**f**) between 1.5km and 3km; Downwind	0.0302	−0.0092	−0.0489
	(0.2739)	(0.1801)	(0.1392)
Tons of pesticide (**g**) between 3km and 25km; Upwind	−0.0270	−0.0210	−0.0966
	(0.1593)	(0.1218)	(0.1246)
Tons of pesticide (**g**) between 3km and 25km; Orthogonal	0.0445	0.1522**	0.1454**
	(0.0625)	(0.0592)	(0.0654)
Tons of pesticide (**i**) between 3km and 25km; Downwind	0.1280**	0.2272***	0.1982***
	(0.0527)	(0.0519)	(0.0609)
Monitor FE	T	T	T
Day-of-sample FE	T	F	F
Week-of-sample FE	F	T	F
Month-of-sample FE	F	F	T
N	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Windsorization* replaces any value *above* the 97.5th percentile of a variable with the 97.5th percentile. Windsorization attempts to limit the influence of extreme values within the independent variables(s). *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table C.4: **Increases in mean PM2.5 from aerial pesticides**: Wind-variation results, Windsorized

Dependent variable: Mean daily PM2.5 level		
	(1)	(2)
Tons of pesticide (**a**) within 1.5km; Upwind	0.2604* (0.1377)	0.2961** (0.1393)
Tons of pesticide (**b**) within 1.5km; Orthogonal	−0.1157 (0.0997)	−0.1079 (0.1018)
Tons of pesticide (**c**) within 1.5km; Downwind	0.2294 (0.1617)	0.2871 (0.1962)
Tons of pesticide (**d**) between 1.5km and 3km; Upwind	0.1421 (0.1117)	0.1659 (0.1306)
Tons of pesticide (**e**) between 1.5km and 3km; Orthogonal	0.1427 (0.1537)	0.1242 (0.1557)
Tons of pesticide (**f**) between 1.5km and 3km; Downwind	−0.2580 (0.2255)	−0.2195 (0.2159)
Tons of pesticide (**g**) between 3km and 25km; Upwind	−0.1045 (0.1421)	−0.1484 (0.1421)
Tons of pesticide (**g**) between 3km and 25km; Orthogonal	0.1078 (0.0741)	0.0917 (0.0750)
Tons of pesticide (**i**) between 3km and 25km; Downwind	0.1746* (0.0927)	0.1545* (0.0843)
Monitor FE	T	T
Week-of-year FE	T	F
Month-of-year FE	F	T
Day-of-week FE	T	T
Year FE	T	T
N	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Windsorization* replaces any value *above* the 97.5th percentile of a variable with the 97.5th percentile. Windsorization attempts to limit the influence of extreme values within the independent variables(s). *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table C.5: **Increases in max. PM2.5 from aerial pesticides**: Wind-variation results, Windsorized

Dependent variable: Maximum daily PM2.5 level	(1)	(2)	(3)
Tons of pesticide (**a**)	0.5264***	0.3499***	0.3399**
within 1.5km; Upwind	(0.1745)	(0.0799)	(0.1334)
Tons of pesticide (**b**)	−0.2352	−0.7615***	−0.3614***
within 1.5km; Orthogonal	(0.2582)	(0.1061)	(0.1333)
Tons of pesticide (**c**)	0.2621**	0.1662**	0.1430
within 1.5km; Downwind	(0.1273)	(0.0724)	(0.2130)
Tons of pesticide (**d**)	0.1283	0.1610	0.2410*
between 1.5km and 3km; Upwind	(0.1817)	(0.1314)	(0.1264)
Tons of pesticide (**e**)	0.0328	0.0672	0.1092
between 1.5km and 3km; Orthogonal	(0.3910)	(0.2236)	(0.1692)
Tons of pesticide (**f**)	0.1200	−0.0378	−0.1210
between 1.5km and 3km; Downwind	(0.3151)	(0.3269)	(0.2205)
Tons of pesticide (**g**)	−0.0838	−0.1033	−0.1820
between 3km and 25km; Upwind	(0.1787)	(0.1373)	(0.1428)
Tons of pesticide (**g**)	0.0583	0.1649**	0.1571*
between 3km and 25km; Orthogonal	(0.0748)	(0.0712)	(0.0808)
Tons of pesticide (**i**)	0.1262***	0.2586***	0.2113***
between 3km and 25km; Downwind	(0.0467)	(0.0611)	(0.0665)
Monitor FE	T	T	T
Day-of-sample FE	T	F	F
Week-of-sample FE	F	T	F
Month-of-sample FE	F	F	T
N	26,242	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Windsorization* replaces any value *above* the 97.5th percentile of a variable with the 97.5th percentile. Windsorization attempts to limit the influence of extreme values within the independent variables(s). *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

Table C.6: **Increases in max. PM2.5 from aerial pesticides**: Wind-variation results, Windsorized

Dependent variable: Maximum daily PM2.5 level	(1)	(2)
Tons of pesticide (**a**) within 1.5km; Upwind	0.3085** (0.1301)	0.3422** (0.1343)
Tons of pesticide (**b**) within 1.5km; Orthogonal	−0.4266*** (0.1188)	−0.3842*** (0.1312)
Tons of pesticide (**c**) within 1.5km; Downwind	0.1274 (0.1493)	0.1756 (0.1828)
Tons of pesticide (**d**) between 1.5km and 3km; Upwind	0.2430* (0.1372)	0.2657* (0.1584)
Tons of pesticide (**e**) between 1.5km and 3km; Orthogonal	0.2582 (0.2262)	0.2714 (0.2353)
Tons of pesticide (**f**) between 1.5km and 3km; Downwind	−0.3153 (0.3397)	−0.2958 (0.3098)
Tons of pesticide (**g**) between 3km and 25km; Upwind	−0.1947 (0.1530)	−0.2469 (0.1559)
Tons of pesticide (**g**) between 3km and 25km; Orthogonal	0.1032 (0.0853)	0.0822 (0.0852)
Tons of pesticide (**i**) between 3km and 25km; Downwind	0.1744** (0.0883)	0.1500* (0.0798)
Monitor FE	T	T
Week-of-year FE	T	F
Month-of-year FE	F	T
Day-of-week FE	T	T
Year FE	T	T
N	26,242	26,242

Notes: Each column denotes a separate regression. Errors are two-way clustered within (1) EPA sensor and (2) day of sample. *FE* refers to *fixed effect*. *Windsorization* replaces any value *above* the 97.5th percentile of a variable with the 97.5th percentile. Windsorization attempts to limit the influence of extreme values within the independent variables(s). *Upwind* refers to the degree to which pesticide applications occured upwind of the EPA monitor—0 and 60 degrees (in absolute value). Similarly, I label angles between 60 and 120 degrees as *Orthogonal* and angles between 120 and 180 degrees as *Downwind*. See the Empirical Strategy section for a detailed explanation. I two-way cluster the errors by (1) monitor and (2) day of sample. *Significance levels:* *10%, **5%, ***1%. The letters in parentheses (*e.g.,* '(**a**)') reference labeled areas in the figures associated with these results.

C.3 Data appendix

EPA monitoring data

I downloaded the EPA PM2.5 monitoring data from the EPA's I use daily data from both FRM/FEM mass systems and non-FRM/FEM mass systems.[1]

PUR data

I only use PUR data related to agricultural production (`record_id` of `2` or `C`) and for pesticides that were applied aerially (`type = A`). Because both empirical strategies rely on geography and timing, I drop use reports missing dates or coordinates (*i.e.*, no `NA`s).

With regards to the California DPR's map of sections, I drop 42 sections (approximately 0.03% of sections) that map to multiple polygons—again because both empirical strategies rely on confidently locating pesticide applications in space. The dropped sections account for approximately 0.2% of the pounds of pesticide reported in the PUR system.

Tables C.1 and C.2 use Windsorized values of the PUR data. Specifically, I Windsorize the variable that represents the total pounds of pesticide applied, replacing any value that exceeds the 97.5$^{\text{th}}$ percentile with the 97.5$^{\text{th}}$ percentile.

Wind data

The NLDAS-2 data come from the Goddard Earth Sciences Data and Information Services Center (GES DISC). While the NLDAS-2 generates hourly data, the paper uses wind estimated at noon at 10 meters above the ground.[2] I calculate wind speed and direction using the U and V wind vectors and trigonometry:

$$\text{Wind}_{\text{Speed}} = \sqrt{u^2 + v^2} \tag{C.1}$$

$$\text{Wind}_{\text{Angle}} = \tan^{-1}\left(\frac{u}{v}\right) \times \frac{180}{\pi} \tag{C.2}$$

In addition, geographic resolution is not the only source of noise in the wind-based measurements. While the NLDAS-2 provides NASA's best attempts to recreate historical wind outcomes at a high spatiotemporal level, the the NLDAS-2 wind data likely introduce additional noise and, consequently, attenuation.

[1] *FRM* is the acronym for Federal Reference Method; *FEM* abbreviates Federal Equivalent Methods.
[2] The time-of-day data in the PURs do not appear to meet data-quality standards.