DEFINING AND MEASURING ENTREPRENEURSHIP FOR REGIONAL RESEARCH: A NEW APPROACH

BY

SARAH A. LOW
B.S., Iowa State University, 2002
M.S., Purdue University, 2004

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Consumer Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Doctoral Committee:

Professor Andrew M. Isserman, Chair
Professor Edward Feser
Professor Geoffrey J. D. Hewings
Professor Randall Westgren, University of Missouri

ABSTRACT

A strong correlation might exist between entrepreneurship and long-term regional employment growth (Acs and Armington, 2003). Entrepreneurship may be a more sustainable economic development strategy than alternatives, like industrial recruitment, because entrepreneurs tend to locate in their home region. Research and policies on fostering entrepreneurship are hindered, however, by the lack of a clear definition and measure of entrepreneurship (Bruyat and Pierre-Andre, 2000). Multiple definitions of entrepreneurship, often flawed, lead to contradictory findings that fuel policymaker confusion (Tamasy, 2006). Most importantly, the commonly used measures of entrepreneurship ignore innovation—a long established defining attribute of entrepreneurship for economic development. This is problematic because only a fraction of new businesses are innovative (Audretch, 2005). Reliable measures of entrepreneurship must be developed to make possible better economic development research and more effective economic development strategies.

In this dissertation, I develop a definition and regional measure of entrepreneurship that will aid entrepreneurship research and economic development policy. I address defining and measuring entrepreneurship, posit a comprehensive definition of entrepreneurship, and develop a method for measuring entrepreneurship that does not ignore the innovation attribute. I test the relationship between economic growth and the new entrepreneurship measures, and estimate the determinants of entrepreneurship using the new measures. The measure I develop is unique, differing from other available measures because it measures the most innovative of entrepreneurs.

Chapter 1 motivates the need for a different regional measure of entrepreneurship. Chapter 2 posits a three-part definition of entrepreneurship, with roots in the work of early entrepreneurship scholars including Schumpeter, Knight, and Say. Chapter 3 reviews current measures of entrepreneurship and compares them to the I present a multi-faceted definition of entrepreneurship and create an annual county-level indicator that incorporates innovation—a commonly overlooked aspect of entrepreneurship. The lack of a clear definition and measure of entrepreneurship hinders the research informing entrepreneurial support policies (Bruyat and Pierre-Andre, 2000). Confusion amongst

policymakers arises from definitions that are either incomplete or contradictory (Tamasy, 2006). Despite measurement problems, entrepreneurial support programs are popular and effective economic development strategies. Since entrepreneurs often locate in their home region, entrepreneurial support may prove to be a more effective economic development strategy than prominent strategies such as industrial recruitment. Stronger economic development research and more effective economic development strategies require more reliable measures of entrepreneurship.

Chapter 4 develops new indicators of entrepreneurship that capture all three components of the proposed definition. The identification of innovative industries, industries with high level of skill, technology, patents, churn, and employment growth, using detailed NAICS (North American Industrial Classification System) industry data, represents an important contribution of this dissertation. By applying the innovative industries to single-unit employer establishment birth and self employment data, I create county-level measures that are available annually for all counties. Using the reduced-form model of entrepreneurship developed by Goetz and Rupasingha (2008), Chapter 5 assesses the determinants of the new entrepreneurship indicator. In Chapter 6, I use a growth model recently developed at the U.S. Department of Agriculture's Economic Research Service (McGranahan, Wojan, and Lambert, 2009) to examine the relationship between my new indicator of entrepreneurship and economic growth. I find a positive and robust relationship between growth and my new indicator of entrepreneurship. Chapter 7 reviews the results and addresses policy-implications, problems, and future work.

My new indicators represent an improvement over current measures of entrepreneurship and have the potential to improve entrepreneurship research and policymaking. The chief contribution of these new measures is that they incorporate innovation, which others ignore. These indicators are imperfect, but nevertheless represent a significant contribution to the literature and can stimulate discussion among entrepreneurship scholars about how we conceptualize and measure entrepreneurship.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help and support of many people. Many thanks to my advisor, Andrew M. Isserman, for reading and revising my work and for teaching me about writing and story telling along the way, for ideas and encouraging me to think outside the box, and for encouraging me to attend the University of Illinois and providing assistantship support. In addition, thanks to my committee members, Edward Feser, Geoffrey J. D. Hewings, and Randall Westgren, who offered guidance and asked interesting questions. Thanks to USDA Economic Research Service for providing support for me to complete this dissertation during the 2008-2009 school year, and especially Mary Ahearn, Mary Bohman, Jim Johnson, David McGranahan, Tim Wojan, and seminar participants. Thanks to the REAP and REAL Mafia at the University of Illinois for moral support, my fellow ACE Ph.D. students, and friends Mallory Rahe, Stephan Weiler, and Bed Wood. I extend a special thanks to Mark C. White for moral support, editing, and formatting assistance, and enduring this long process with me (or more appropriately, enduring me during this long process). Thanks to my late mother, Sharon Low, who was editing this dissertation in a hospital bed until her death and whose spirit gave me the energy to complete my dissertation. Finally, thanks to my grandparents, Otto and Delores Stender, John, Dee Dee, and Kennedy Low, and the rest of my family. Also a special thank you to my dog, Stewart, for his unconditional love and support over the past ten years.

To My Mother, Sharon L. (Stender) Low

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Policies and programs to foster entrepreneurship, particularly at the state and local level, are becoming increasingly common. The lack of a theoretically sound definition and appropriate measure of entrepreneurship, however, hinders effective policymaking and research. Existing research uses a multitude of entrepreneurship indicators, each identifying one or more attributes of entrepreneurship, each partially dictated by the availability of data for the region and time period of interest. Most important, the commonly used measures of entrepreneurship ignore innovation—a long established defining attribute of entrepreneurship for economic development. Researchers must development more reliable measures of entrepreneurship in order to strengthen economic development research and create more effective economic development strategies.

This dissertation presents a conceptually clear definition of entrepreneurship and indicators of this definition for use in economic development research and policymaking. A key aspect of these indicators is that they capture innovation better than existing measures of entrepreneurship. This dissertation contributes a method for identifying *innovative industries* and the *Entrepreneurial Industries* entrepreneurship indicators, which have both the breadth and depth to be useful for regional research and economic development purposes.

My new indicators represent an improvement over current measures of entrepreneurship and have the potential to improve research and policy by improving the quality of empirical entrepreneurship research. The indicators are interesting because they are the first known attempt to create an indicator that captures multiple facets of entrepreneurship and are readily available at the county level. Although my indicators are imperfect, this research represents a significant contribution to the literature and I hope it stimulates discussion among entrepreneurship scholars about how we measure and conceptualize entrepreneurship.

In Chapter 2, I discuss functional definitions of entrepreneurship, paying particular attention to how definitions of entrepreneurship relate to economic development. I identify three broad attributes in existing definitions of entrepreneurship and posit a definition for this dissertation that includes the three attributes. I define entrepreneurship as 1) owning or operating a firm to capture economic rents, while 2) bearing the risk and uncertainty of the firm, and 3) being innovative or continually reallocating resources.

In Chapter 3, I compare existing measures of entrepreneurship and discuss how they relate to the definition this dissertation uses. No existing measure meets the definition of entrepreneurship established in Chapter 2, with innovation being the most overlooked attribute of entrepreneurship.

In Chapter 4, I respond to the call for the development of regional entrepreneurship measures that capture the innovative nature of entrepreneurship better than existing measures. I identify innovative industries using occupation skill and technology, and industry patenting, churn, and employment growth. I use data on single-unit employer establishment births and self employment to count establishments in innovative industries for each county. The establishment birth data are available at the five-digit NAICS (North American Industrial Classification System) industry level for U.S. counties, annually. These data from the Dynamic Data, U.S. Statistics of Business and were obtained by USDA-ERS through a special agreement with the Census Bureau. The self employment data are available annually for U.S. counties from the Census Bureau's Nonemployer Statistics series. These data are available at the six-digit NAICS industry level, but because the data are publicly available, they are subject to suppression.

In Chapter 5, I examine determinants of entrepreneurship and the determinants of my new indicator using an empirical model of county-level entrepreneurship developed by Goetz and Rupasingha (2008). I find the determinants of the new indicator are similar to parent measures, but amenities, urbanization, and financial collateral appear to drive Entrepreneurial Industries.

In Chapter 6, I test the new entrepreneurship indicators in a growth model recently developed at the USDA Economic Research Service. I test the relationship between the new indicators and employment, population, and job growth. I find a robust, positive relationship between Entrepreneurial Industries and growth, which may be stronger than the relationship other measures have with growth, likely because the Entrepreneurial Industries indicator includes the most innovative establishments.

Chapter 7 offers conclusions and discussion on the virtues and vices of the Entrepreneurial Industries indicator. I also discuss the dissertation's other research contributions and conclude with what I learned during the dissertation process.

CHAPTER 2: DEFINING ENTREPRENEURSHIP

"risk-takers, the doers, the makers of things"
--President Obama on entrepreneurs,
Inauguration Day, 2009

The entrepreneur has played an important role in the academic literature for 250 years. While there remains a broad consensus about the central role of entrepreneurship in the economy, theoretical and conceptual models of entrepreneurship vary widely. Theoretical models of entrepreneurship are weak or non-existent, and the term entrepreneur is still vaguely defined, even though entrepreneurship scholars seem obsessed with defining the word entrepreneur (Bull and Willard, 1993). Scholars have long disagreed about the definition of entrepreneurship (Cole, 1942). Defining entrepreneurship and developing a theoretical model present two related problems--defining entrepreneurship is hindered by difficulties in conceptualizing and quantifying theoretical models of the entrepreneurial process (Iversen et al., 2008), while the lack of a consensus definition hinders theoretical model development (Bull and Willard, 1993). No theory of entrepreneurship has been developed that explains or predicts when an entrepreneur, by any definition, might appear or engage in entrepreneurship (Bull and Willard, 1993). Many different functional definitions or theories of entrepreneurship have been proffered, likely because entrepreneurship is a dynamic and complex phenomenon with multiple purposes (Bruyat and Pierre-Andre, 2000). This complexity makes it impossible to capture the totality of entrepreneurship without using a multi-component definition (Iversen et al., 2008).

Despite the lack of a consensus definition of entrepreneurship (Iversen et al., 2008; Bull and Willard, 1993; Bruyat and Pierre-Andre, 2000), and confusion in measuring entrepreneurship (Gartner and Shane, 1995; Luger and Koo, 2005; Hoffmann et al., 2006), research on entrepreneurship for economic development is booming. Researchers have found a strong correlation between entrepreneurship and long-term regional employment growth (Acs and Armington, 2003). This relationship has important policy implications as entrepreneurship is often considered a more sustainable economic development strategy than alternatives such as industrial recruitment. Nevertheless, the lack of a theoretically sound definition of entrepreneurship precludes a full understanding of the regional development opportunities associated with entrepreneurship (Casson, 2003).

Good science must begin with good definitions (Bygrave and Hofer, 1991), and in this regard, current entrepreneurship research fails due to definitional ambiguity. We need a clear definition of entrepreneurship to advance theoretical and empirical research that can better inform economic development professionals and policymakers about how entrepreneurship can drive economic development.

This chapter presents a conceptually clear working definition of entrepreneurship for economic development. This definition is based on a review of others' functional definitions of entrepreneurship. I use this definition as the basis for developing new measures of entrepreneurship in Chapter 4.

2.1. THREE ATTRIBUTES OF ENTREPRENEURSHIP FROM THE FUNCTIONAL ENTREPRENEURSHIP LITERATURE

As the theory behind, and definition of, economic entrepreneurship develops, the functions of entrepreneurs receives more attention (Casson, 2003). Literature is moving away from the supply-side (trait-based) approach to defining entrepreneurship, e.g., Low and Macmillan (1988), to a more demand-side approach. The demand-side approach defines entrepreneurship by the entrepreneur's function, or what entrepreneurs do, rather than who entrepreneurs are, and this proves more useful for prescriptive policy research (Gartner, 1990; Rocha and Birkinshaw, 2007). This section discusses several major functional definitions of entrepreneurship used in the economic and economic growth literature over the past 250 years. Particular attention is paid to the relationship between entrepreneurship and economic development. Many definitions of entrepreneurship exist, but the literature points to three broad yet distinct attributes of the entrepreneur's function:

1. Ownership or operation of a firm,
2. Risk and uncertainty bearing, and
3. Innovation or the reallocation of resources.

This section is organized around these three attributes of entrepreneurship.

2.1.1 Ownership or Operation of a Firm

Ownership or operation of a firm is an important attribute of entrepreneurship. It is not sufficient to define entrepreneurship, but I posit it is necessary to define entrepreneurship. The exploitation of entrepreneurial ideas must take place within a firm, as there is no market for entrepreneurship (Casson, 2003; Ross and Westgren, 2006). As a result, owning or operating a firm—particularly a small firm—is one of the most widely used definitions of entrepreneurship (Georgellis and Wall, 2000; Parker, 1996; Glaeser, 2007; Goetz and Rupasingha, 2008; Shrestha et al., 2007). The owner or operator of a firm is the firm's leader. The firm leader makes daily business decisions about innovation, risk preferences, and coordinates firm activities (Cantillon, [1755] 1964; Casson, 2003). As will be shown, numerous theorists use ownership of a firm as one of the key elements in defining entrepreneurship.

Richard Cantillon (1680-1734), an Irish economist, was the first economist to define entrepreneurs by their function (Rocha and Birkinshaw, 2007). Cantillon's entrepreneur is a firm operator, who has an ownership stake but also bears risk. Cantillon's entrepreneur differs from a capitalist because he/she directs production and his/her function is to equate supply and demand in the market. By contrast, a capitalist simply provides capital and does not operate the firm (Cantillon, [1755] 1964).

Jean-Baptiste Say (1767-1832) also distinguishes the entrepreneur from capitalists and laborers, but Say defines the entrepreneur as a manager. Say's entrepreneurs are a factor of production whose job it is to assess firm opportunities and select the most favorable (Say, [1803] 2001). Say affirms that the entrepreneur receives a wage premium due to the scarcity of his/her skills, akin to Coase's Theory of the Firm (1937). Say does not emphasize the risk bearing nature of entrepreneurship like Cantillon does, but instead focuses on the managerial, or operator functions. These functions include combining factors of production in the firm in the most efficient manner (Iversen et al., 2008).

T.W. Schultz (1902-1998) was an agricultural economist in the Chicago school, and his main contribution was the human capital theory of entrepreneurship. Schultz defines entrepreneurship as the ability to reallocate efficiently resources to deal with disequilibria in the market and maximize profit. These are decisions that an owner or operator must make (Klein and Cook, 2006; Iversen et al., 2008). Schultz posits that economic growth comes from individuals responding to disequilibria, and the higher their human capital, the more optimal are their responses to changing economic conditions. Schultz extends the entrepreneurship theory literature by concluding that economic growth can be advanced with entrepreneurs who have high levels of human capital.

Like Cantillon, Say, and Schultz, Mark Casson (1945-) defines entrepreneurs by their operator function—assessing markets, making decisions, negotiating, and coordinating firm activities. Casson (2003) differentiates the entrepreneur and manager, however, by positing that the entrepreneur establishes a firm and bears the start-up costs necessary to exploit his entrepreneurial behavior and pursue profit. Casson's entrepreneurs specialize in decision-making, but Casson also makes clear that not all decision makers are entrepreneurs (Iversen et al., 2008).

2.1.2 Risk or Uncertainty Bearing

Risk and uncertainty bearing are important attributes of entrepreneurship because they distinguish entrepreneurs from wage and salary workers (Knight, 1942; Casson, 2003). Entrepreneurs may be richly rewarded with rents due to innovation and early adoption, but, to be rewarded, they must bear the associated risk and uncertainty. Moreover, risk bearers retain only net profits, after outstanding obligations are paid. Von Thünen, Knight, Cantillon, and Casson all emphasize that the entrepreneur bears the cost of establishing a firm, receives uncertain compensation, and has a low level of risk or uncertainty aversion.

Johann Heinrich von Thünen (1783-1850) worked on marginal productivity and defines economic rents as those that are earned at the margin of production and are created by spatial variation (von Thünen, [1826] 1960). Like Cantillon, von Thünen's entrepreneur bears risk and uncertainty, receiving only the residual profits after he/she makes all contractual payments, but von Thünen was the first to distinguish between risks that can be insured and uncertainty that cannot (Cantillon, [1755] 1964).

Von Thünen and Cantillon's work served as a foundation for the work of Frank Knight (1885-1972), who fleshes out the unpredictable entrepreneurial income component, distinguishing risk from uncertainty in his famous dissertation, "Risk, Uncertainty, and Profit" (Knight, 1942). The association of entrepreneurship with uncertainty provided the early foundation for the American or Chicago School of economic theory. Knight defines the entrepreneur as a firm owner who purchases inputs (labor, raw materials) for a fixed price and makes a product or service, and due to changing preferences, will receive an uncertain price in an uncertain economy. Knight's entrepreneur bears the cost of innovation. Since unpredictable contingencies occur, innovation must be associated with risk-taking and judgment (better conception of the unknowable future market).

Knight (1942) argues that the entrepreneur assumes three functions or tasks:

1. Initiate innovations or useful changes,
2. Adapt to changes in the economic environment, and
3. Assume the consequences of uncertainty related to the innovation.

Knight states that the entrepreneur functions as an economic pioneer by initiating innovations and bearing the costs associated with the innovation's risk and uncertainty. For bearing firm risk and uncertainty, the entrepreneur is entitled to residual income after all contractual payments have been made (Casson, 2003). The innovator is generally more dynamic than the manager who performs routine activities. Knight, does however admit that managers of large firms must make predictions—much like entrepreneurs—but the manager is not the sole recipient of net profits.

Knight defines risk as randomness with a known ex-ante probability distribution, while uncertainty is randomness with an undefined probability distribution (Klein and Cook, 2006). Uncertainty is one of the problems associated with developing a theoretical model of entrepreneurship, because of the unknown probability distribution. Knight argues that entrepreneurs have an unusually low level of uncertainty aversion (Baumol, 1993).

2.1.3 Innovation

Scholars such as Schultz, Kirzner, Knight, and Schumpeter incorporate innovation into their definitions of entrepreneurship. Innovation is a crucial component of entrepreneurship because it is closely connected with the ability to deal with market disequilibria. Many theoretical definitions of entrepreneurship incorporate initiating innovation (Schumpeter and Opie, 1983) and/or recognizing market opportunities (Schultz, 1975).

Two scholars, Schultz and Kirzner, write that market opportunities and reallocating resources in response to these market opportunities is entrepreneurship, not initiating innovation. Schultz defines entrepreneurship as efficiently reallocating resources and dealing with disequilibria in the market to maximize profit (Klein and Cook, 2006; Iversen et al., 2008). Schultz (1975) argues that disequilibria exist, not because the entrepreneur does not see them, but because reallocating resources takes time. Israel Kirzner does not view returns to entrepreneurship as compensation for uncertainty (Ross and Westgren, 2006), but rather defines entrepreneurs as those who recognize profit opportunities brought about by economic shocks and move the economy towards equilibrium (Baumol, 1993).

Unlike Kirzner and Schultz, Knight and Schumpeter's entrepreneur creates disequilibrium in the market economy that necessitates innovation or change (Knight, 1942; Schumpeter and Opie, 1983). He/she is responsible for initiating and adapting to economic changes and capturing scarce monopoly rents until those rents fall to zero. Knight's entrepreneur shocks the economy with innovation and as those innovations become adopted and diffused, he/she adapts to the changing market. Schumpeter's entrepreneur, however, is complex and worthy its own discussion.

Joseph Schumpeter (1883-1950) represents the German school of economics that emphasizes entrepreneurship and innovation. Schumpeter believes the entrepreneur is the innovator who transforms inventions and ideas into economically viable entities (Baumol, 1990). Schumpeter defines the entrepreneur as someone motivated by profit to destroy outdated patterns of thought and action. Notably, Schumpeter did not think of all businessmen or capitalists as entrepreneurs because the entrepreneur can obtain credit, thereby making capital unnecessary.

Schumpeter is widely known for his definition of creative destruction—the start-up of new firms and displacement of the incumbents, thereby establishing superior economic performance in terms of both innovation and growth (Schumpeter and Opie, 1983). Schumpeter argued that innovation was the strategic stimulus for economic development; thus, innovation is a natural component of any definition of entrepreneurship for economic development (Schumpeter and Opie, 1983). Innovation was the lynchpin of economic development to Schumpeter. Schumpeter lays out five tasks that lead to innovation (McGraw, 2007, Schumpeter and Opie, 1983).

1. Introduction of a new good, or a new quality of good
2. Introduction of a new method of production
3. The opening of a new market
4. The conquest of a new source of supply of raw materials or half-manufactured goods
5. The carrying out of the new organization

These tasks suggest that Schumpeter thought of innovation as multi-faceted and included product, process, organization, purchasing, and marketing innovations. Including innovation in my definition of entrepreneurship allows for a qualitatively different measure of entrepreneurship, by enabling me to capture entrepreneurs who both create products and processes, rather than simply operate a small business.

Despite Kirzner, Schultz, Knight, and Schumpeter incorporating innovation into their theoretical definitions of entrepreneurship, most empirical definitions of entrepreneurship overlook innovation, principally because innovation is difficult to measure. Thus, as of this writing, only second-best measures of innovation are available (Green et al., 2006).

2.1.4 Comparing Definitions of Entrepreneurship

Table 2.1 presents these three widely recognized attributes of entrepreneurship—owner/operator, risk/uncertainty bearing, and innovation—and how definitions of entrepreneurship consider these three attributes. Even though these are commonly recognized attributes, Table 2.1 shows that no one functional definition adequately incorporates all three attributes of entrepreneurship.

Table 2.1 Comparison of Definitions of Entrepreneurship

	Owner/Operator		Risk Bearer		Innovator	
	ownership	operation	risk	uncertainty	new combination	reallocation
Cantillon	X	X	X			
von Thünen				X		
Say		X				
Schumpeter		X			X	
Knight			X	X	X	X
Kirzner						X
Schultz		X				X
Casson	X	X	X	X		

2.2 PROPOSED DEFINITION OF ENTREPRENEURSHIP

The complexity of entrepreneurship makes it impossible to capture the totality of entrepreneurship with one idea; therefore, I propose the following definition of entrepreneurship:

The entrepreneur has an owner or operator function, a risk and uncertainty bearing function, and, perhaps most importantly, an innovation function.

The combination of innovation, owning or operating an establishment, and bearing risk/uncertainty provides an effective working definition of entrepreneurship that is useful for economic development purposes. This definition captures all of the components other scholars identify and is multi-faceted, to capture these multiple components of entrepreneurship (Figure 2.1).

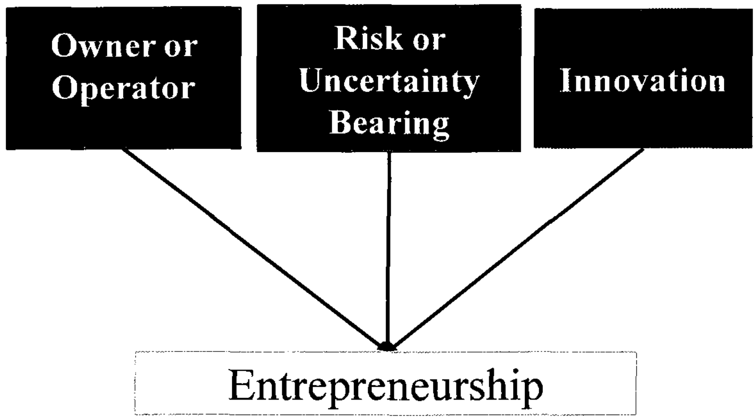

Figure 2.1 Multi-Faceted Definition of Entrepreneurship

The owner or operator function differentiates entrepreneurs from intrapreneurs and social entrepreneurs, by ensuring that the entrepreneur has a firm within which he/she can capture rents and capitalize on entrepreneurial skills. Risk bearers are the residual claimant to rents and face uncertain profits because employees and creditors must be paid first, leaving a positive or negative residual for the risk bearing entrepreneur. This uncertain return stimulates entrepreneurs who hope the return is lucrative. Finally, innovators create novel combinations of goods, services, and markets in response to economic opportunities, differentiating themselves from small business owners who do not innovate. Diminishing rents motivate entrepreneurs to constantly innovate and reallocate resources to capture changing market opportunities.

This definition of entrepreneurship, like others, is difficult to formalize in a mathematical model. Kirzner argues that the entrepreneur is inherently unpredictable—making a predictive theory of entrepreneurship impossible (Casson, 2003). A formal mathematical or theoretical model has been the goal of many economists studying entrepreneurship, but to date none has been widely accepted. The inadequacy of economic theory in explaining dynamic processes and heterogeneous firms' actions in a solvable model has been the greatest hindrances to the development of a widely accepted model. Neoclassical models are easier to derive, but homogeneous firm and zero profit assumptions combined with the lack of dynamic modeling diminishes this approach. Parker (2006) uses optimal control theory to develop a dynamic model that optimizes individual behavior, but is still limited by neoclassical assumptions. Endogenous growth theory removes the zero profit assumption but remains static and "entrepreneur-less," because firms are homogenous. Developing a theory of economic dynamics will be crucial for the advancement of economic theory, but could also prove very useful for research on both economic development and entrepreneurship (North, 1994).

2.3 CONCLUSION

In this chapter, I have established a three-part conceptual definition of entrepreneurship, capturing the principal components of many functional theories of entrepreneurship. Although my definition may be imperfect, good science must begin with a good definition. This definition will serve the dissertation's purposes of contributing to the entrepreneurship and economic development literature, stimulating discussion among scholars about how entrepreneurship is conceptualized and measured, and providing a theoretically sound definition of regional entrepreneurship.

CHAPTER 3: CURRENT MEASURES OF ENTREPRENEURSHIP

Measures of entrepreneurship utilized in economic development research and policymaking are based, not on ideal definitions of entrepreneurship, but on what data are available—a class of "second best" measures (Green et al., 2006). Many studies focus on the measurement of entrepreneurship (Gartner and Shane, 1995; Luger and Koo, 2005; Hoffmann et al., 2006) but no measure is clearly superior to others. Each metric has its own strengths and weaknesses and the choice of measure is likely to influence the research results (Gartner and Shane, 1995).

In this chapter, I discuss the strengths and weaknesses of existing entrepreneurship measures, categorizing them by self employed, establishments, and births, and comparing each to the definition of entrepreneurship. I find that commonly used measures of entrepreneurship 1) ignore innovation, because it is difficult to quantify, and 2) are data-driven rather than driven by theory or definition.

3.1 SELF EMPLOYMENT

The self employment rate is the most widely used measure of entrepreneurship in economic development applications and regional research (Iversen et al., 2008). Researchers have recognized self employment as a seedbed of entrepreneurship—and a convenient measure of entrepreneurs in a region (Low et al., 2005; Goetz and Rupasingha, 2008). Self employment is a stock measure, taken at one point in time, and stock measures are more stable year-to-year than flow measures, thus more suitable for cross-sectional studies (Gartner and Shane, 1995). Despite wide use of the self employment rate as a measure of entrepreneurship, it is an imperfect measure because it is very broad and captures all types of small business activity, not necessarily entrepreneurial activity (Acs et al., 2008).

In this section, I discuss my analysis of self employment in the regional entrepreneurship literature. I find the self employment rate is an imperfect measure of jobs held by those who work for themselves; it is easy to obtain and captures entrepreneurial activity but not the extent to which entrepreneurs are successful or innovative (Munn, 2008). The self employment rate does meet owner/operator and risk-bearing attributes of entrepreneurship but does not meet the innovation attribute. Ideally, we could measure activity of the innovative self employed—those offering new services, innovative products, or unique methods of production or delivery. Users should recognize that self employment is just a measure of self employment, not a definition of entrepreneurship (Georgellis and Wall, 2000).

3.1.1 The Use of Self Employment as a Measure of Entrepreneurship

The self employment rate is often used to measure entrepreneurship because of its simplicity and availability. The self employment rate has been used for country-level studies (Iversen et al., 2008; Blanchflower, 2004; Blanchflower, 2000; OECD 2000; Parker 2005); regional studies (Parker, 1996; Georgellis and Wall, 2000; Shrestha et al., 2007; Glaeser, 2007 ; Goetz and Rupasingha, 2008; Acs et al., 2008; Saxenian, 1994; Schiller and Crewson, 1997); and longitudinal and panel studies of individual behavior (Baumol, 1993; Lazear, 2005; Reynolds and Curtin, 2008; Hamilton, 2000; Tamasy, 2006; Blanchflower and Oswald, 1998). The wide use of the self employment rate is likely because it is easy to measure with administrative records and publicly available data based on administrative records, e.g., the Bureau of Economic Analysis' Regional Economic Information System (BEA-REIS) or the Census Bureau's Nonemployer Statistics in the U.S. The most widely used measure of U.S. county-level self employment is defined as nonfarm proprietors in a county over total nonfarm employment (Low et al., 2005, Henderson et al., 2006, Goetz and Rupasingha, 2008). Researchers have also measured self employment with surveys of individuals, although this is not practical for U.S. counties (Lazear, 2005; Tamasy, 2006; Blanchflower and Oswald, 1998; Baumol, 1993).

3.1.2 Calculating the Self Employment Rate

The self employment rate is more useful for interregional comparisons than the level of self employment. Self employment is usually normalized by employment, rather than population, because workers more closely represent the pool of nascent entrepreneurs. Employment and labor force are commonly used denominators, but U.S. county studies generally use nonfarm employment because nonfarm labor force data are not available for counties. Iversen et al. (2008) show that when calculating the self employment rate, the choice of denominator, labor force vs. employment, can affect the measure. Nobody has examined the choice of denominator for U.S. counties. Thus, I compute the self employment rate for U.S. counties using both denominators, labor force (Bureau of Labor Statistics) and total employment (BEA-REIS), with BEA proprietor employment as the numerator.[1] I do not detect a substantial difference. I find relatively little difference in the mean and spread, the correlation between the two measures is 0.93, and the Spearman Test of Independence rejects the null hypothesis that the two measures are independent of each other.

Most studies exclude farm self employment from the self employment rate because farming is influenced heavily by subsidies (Iversen et al., 2008), there is a relatively high proportion of unpaid family labor in farming operations, and there is wide disparity in regional levels of farm self employment (Meager, 1992; Blanchflower, 2000). The U.S. self employment rate falls by 10 percent when agricultural self employment is excluded (Iversen et al., 2008) and the agricultural self employment rate varies widely across U.S. counties—as high as 79.1 percent and as low as zero percent, with a mean of 20.0 percent.[2] Heterogeneity suggests farm self employment should not be used for regional research. Indeed, Census Bureau data products, e.g., County *Business Patterns,* Nonemployer Statistics, and Statistics of U.S. Businesses, exclude crop and animal production.

3.1.3 Relating Self Employment to Theory

The self employment rate meets two of the three dimensions of entrepreneurship, failing only innovation. Firm ownership or operation and risk and uncertainty bearing are inherent in being self employed, (Noteboom, 1999; Baumol, 1993) but the self employment rate includes many who are not innovators. Only ten to twenty percent of the self employed are innovative (Noteboom, 1999). Publicly available self employment data are not sufficiently refined to measure the activity of innovative self employed

[1] All data are 2006, U.S. counties

[2] Authors own calculation. Calculated as: (proprietor employment –nonfarm proprietor employment)/proprietor employment, 2006.

individuals who offer new services, innovative products or technologies, or unique methods of production or delivery.

The self employed entrepreneur identifies an opportunity, creates an institution to capture the rents associated with that opportunity, and profits from his/her work. These functions are somewhat related to Schumpeterian entrepreneurship, indeed, Schiller and Crewson (1997) posit that the self employment rate is a reasonable proxy for Schumpeterian entrepreneurship, arguing that self employment is a pragmatic, if not compelling, measure of entrepreneurial activity.

I believe the self employment rate does measure Schumpeterian entrepreneurship but it also captures lots of non-Schumpeterian entrepreneurs—those who have not innovated, developed a new product, service, or technology, or, those who have stopped innovating (Georgellis and Wall, 2000). Thus, self employment does not meet by innovation attribute of entrepreneurship. Businesses may start-out fitting the definition of Schumpeterian entrepreneurship, but they rarely remain in such a category (Schumpeter and Opie, 1983) because they stop innovating once established.

The self employed clearly fit Say's definition of entrepreneurship, the owner-manager. Small firm's owners conduct day-to-day tasks associated with running a firm—managing, bookkeeping, marketing, taking out the trash, etc. In this role, the self employed are managers as well as owners. Even the self employed with no employees must assume some managerial roles, as there is no one else authorized to make decisions.

Few argue that the self employed bear risk and uncertainty. Knight posits that the entrepreneur faces risk and uncertainty in his remuneration, an attribute that the self employed hold. Knight's definition of entrepreneurship fits into self employment because most self employment data do not include incorporated establishments. Those who take on the risk of starting a business are more entrepreneurial than wage and salary workers are, whether or not the business is innovative just by risk-taking.

Self employment is widely used because it is readily available to the public, can be calculated for small areal units, and is particularly well suited for regional research due to its availability over space and time. Finally, self employment captures the stock of owner/operators and individuals bearing risk and uncertainty.

3.1.4 Problems with Self Employment as a Measure of Entrepreneurship

Data issues cause the most significant problems with self employment. These issues arise from three primary problems:

1. Self employment requires careful interpretation because it varies greatly across space and time.
2. self employed business owners with employees are excluded from self employment data in the U.S.
3. Part-time or multiple job holding self employed are counted as equal to full-time self employment because data are based on tax returns and there is no information on hours worked or percent of income from self employment.

The self employment rate varies greatly across time and space. Nonfarm self employment rates are higher in nonmetro counties than metropolitan counties and highest in the Great Plains, Southern Appalachia, and parts of the Rockies (Low et al., 2005). Some places that seem especially entrepreneurial actually have low self employment rates, e.g., San Jose (Silicon Valley). Finally, self employment rises in recessionary periods everywhere and rises, in metro areas in especially good economic times (Parker, 1996). Conducting interregional and time-dependent analysis requires researchers to consider this variation across space and over business cycles.

U.S. federal data sources use a narrow definition of self employment that excludes firms with paid employees.[3] This narrow definition of self-employed is also used in Australia and Japan. A broader self employment definition is used throughout Europe and by the Organisation for Economic Cooperation and Development (OECD). Using the OECD definition of self employment, the U.S. self employment rates would be up to 50% higher than reported (OECD, 2000). Summing the number of self-employed and sole owners of corporations and businesses solves this problem, e.g., de Wit (1993), but data on the number of sole owners of corporations and businesses are not available for regions in the United States.

Another difference between the U.S. and OECD data is the treatment of multiple jobholders. U.S. data count jobs rather than individuals but OECD Labor Force Statistics count the *main* job, creating a difference where many self employed are also wage earners, e.g., rural areas where many hold multiple off-farm jobs and/or non-farm

[3] Self employment is calculated using filings of federal tax Form 1040 (Schedule C), for sole proprietorships, and Form 1065 for partnerships.

proprietorships. A solution is to access individuals' tax returns and estimate the percent of personal income that comes from self employment but inquiries Stephan Goetz made to CES and BEA about answering this question with tax return data cited concern about privacy and disclosure of records. Finally, the self employment rate counts all self employed equally, the necessity-driven self employed are equals to the wealthiest and most innovative entrepreneurs (Glaeser, 2007), thus, the number of self employed does not equate to their value to the regional economy.

3.2 ESTABLISHMENT MEASURES OF ENTREPRENEURSHIP

The establishment rate may be a good indicator of past entrepreneurship (Gartner and Shane, 1995; Loveridge and Nizalov, 2006). Chinitz (1961) describes the entrepreneurial culture of New York City, a culture that encourages entrepreneurship and has many self employed and small family businesses, and Pittsburgh, an industrial culture where labor force participants rely upon getting a job at U.S. Steel—one firm with many employees. In Chinitz's example, establishments per capita are high in New York City compared to Pittsburgh. Thus, a high establishment rate is indicative of entrepreneurial climate and is suited for longitudinal entrepreneurship research due to its availability and stability over time (Saxenian, 1994).

I find many disadvantages associated with using the establishment rate to measure entrepreneurship, however. Establishments fail both the innovation and risk/uncertainty attributes of entrepreneurship, making it a weak substitute for entrepreneurship, rather a proxy of past entrepreneurship. Additionally, I find that the establishment rate is spatially dependent; it is high in sparsely populated areas due to market structure, thus care must be taken when using the establishment rate for regions.

3.2.1 Use of Establishment Measures

Gartner and Shane (1995) present *Organizations per capita*, as a measure of entrepreneurship based on the premise of entrepreneurship being ownership. Gartner and Shane argue the number of establishments, normalized by population, is a good measure of regional entrepreneurship over time. The measure is easy to compute and data are easy to obtain for regions and countries over relatively long periods. Gartner and Shane (1995) use population in the denominator of the establishment rate because the population is the pool of consumers.

Average firm size, or the average number of employees, assumes that many small firms are more entrepreneurial large firms (Glaeser, 2007; Saxenian, 1994; Chinitz, 1961; McGranahan et al., 2009). Glaeser (2007) writes that when the same numbers of employees are spread over more firms, there must be more entrepreneurs, or firm leaders, per worker; thus, average firm size is a similar measure to organizations per capita.

3.2.2 Relate to Theory

Establishment measures fail the innovation and risk/uncertainty attributes of my entrepreneurship definition. Establishment based measures of entrepreneurship fail to meet risk/uncertainty because they overestimate the risk-bearing or Knightian function of entrepreneurship. Knight's entrepreneur is the residual claimant to firm profits, so privately held single unit establishments have a "Knightian entrepreneur" somewhere, however, publically held establishments do not meet Knight's entrepreneurial function unless manager compensation is tied to performance.

The establishment rate does not capture innovation due to the coarseness of the measure. Establishment rates suffer from the same problem that self employment rates do—they include many repetitive and non-innovative establishments and are not refined enough to capture innovation or Schumpeterian entrepreneurship.

Establishment-based measures of entrepreneurship meet the owner or operator requirement of the definition of entrepreneurship posited in Chapter 2. The establishment rate measures Say's entrepreneur, the number of "managers" meeting the ownership/operation attribute. More firms equates to more managers and more firm founders, who could be entrepreneurs.

3.2.3 Advantages of Establishment Measures of Entrepreneurship

Advantages of establishment measures may or may not outweigh that establishment measures ignore innovation. Establishment measures are readily available, easy to compute across time and space, and are relatively stable across time—unlike self employment (Gartner and Shane, 1995). These factors make establishment rates one of the best longitudinal measures of regional entrepreneurship (Gartner and Shane, 1995; Acs et al., 2008)—if you consider establishments entrepreneurial.

3.2.4 Disadvantages Associated with Establishment Measures of Entrepreneurship

The principal problem with using establishment measures is that it ignores innovation and risk/uncertainty attributes of entrepreneurship, but measurement issues also exist. Measuring entrepreneurship with establishments per capita assume the ratio of establishments to entrepreneurs remains constant, i.e., if five people jointly found a firm, they are only counted as one establishment, rather than five entrepreneurs. This problem is inherent in establishment-level, rather than individual-level measures. Another downside of using average firm size is that it can be seen as a measure of competitiveness or firm age, but Acs et al., (2008) note these limitations do not preclude it from capture some part of what can be considered to be entrepreneurship. Finally, establishments, even normalized, are dependent on the population density. Glaeser finds average firm size for metropolitan statistical areas is similar across urban areas (Glaeser, 2007), however, I find that the measure varies systematically across the rural-urban continuum; rural counties have a smaller average firm size than urban counties and many more establishments per capita.

3.3 DYNAMIC DATA

The flow of establishments, their births and deaths, represents an alternative to stock measures, e.g., self employment, and enables researchers to measure the *creative destruction* within an economy. Dynamic data, however, are more difficult to obtain than stock data and ignore existing establishments. In this section, I discuss dynamic establishment data and measuring entrepreneurship using these data. I relate the measure to the proposed definition of entrepreneurship and discuss the pros and cons of measuring entrepreneurship with dynamic data.

3.3.1 Prior Use of Dynamic Data Measures of Entrepreneurship

Dynamic data are increasingly being used in the entrepreneurship literature because births and deaths are considered more entrepreneurial than self employment and other traditional measures of entrepreneurship (Acs and Armington, 2003; Lee et al., 2004; Luger and Koo, 2005; Acs and Mueller, 2007; Mueller, 2007). Birch (1981) was the first to study establishment dynamics after he compiled the first micro dataset on U.S. establishments and their dynamics in the 1980s (Acs and Mueller, 2007). Today, better micro datasets are available, such as the Census' Longitudinal Business Database, and one publically available dynamic dataset exists, Dynamic Data from, a subset of the Statistics of U.S. Businesses.

Dynamic data include establishment flows over a period, generally a year, and includes births, deaths, churn, and even survival of employer establishments. The establishment birth rate is the most widely used dynamic measure, it is normalized by employment (Mueller, 2007), population (Lee et al., 2004), or establishments (Reynolds et al., 1994), and used to measure the entry or creation of firms. The death rate measures firms made obsolete, however, few researchers use the exit rate alone, rather they use the "churn rate," the sum of the birth rate plus the death rate. If one has access to micro data, the survival rate of establishments can be calculated, which is superior to gross entry and exit (Acs et al., 2006).

Dynamic data—a flow measure—capture change over a period of time, and better capture Schumpeterian and Kirznerian entrepreneurship because flows can measure entrepreneurship dynamically (Iversen et al., 2008). Flow measures relate less to business ownership rates than the stock measures and, as a result, are better able to capture innovation and reallocation of resources. Thus, dynamic data capture innovation better than stock measures, but flow measures are more difficult and costly to obtain, e.g., a survey of individuals or proprietary data.

3.3.2 Relationship to Entrepreneurship Theory

Publically available dynamic data meet owner/operator and risk/uncertainty attributes of entrepreneurship, and are more likely to meet innovation than stock measures. By assuming an individual is responsible for the birth of an establishment, I can infer that new establishments meet Cantillon's entrepreneurship function, ownership or operation of a firm, and Knight's risk and uncertainty bearing because the firm owner bears associated risks. Multi-unit firms/establishments and those organized as corporations are more likely to fail owner/operator and risk/uncertainty attributes of entrepreneurship than single-unit establishments; indeed most establishments begin as S-corporation or partnership, rather than a C-corp.

Schumpeter wrote that births and deaths are essential for innovation, entrepreneurship, and economic growth, indeed, births and deaths capture the essence of Schumpeter's *creative destruction.* Births and deaths, however, do not imply innovation. Dynamic data cannot measure innovation in process, product, or markets. Reynolds et al. (1994) find most single-unit establishment births are replicative, making them unsuitable for capturing innovation. Dynamic data capture more innovation than stock data, births and deaths do not meet the innovation attribute of entrepreneurship because so many new establishments are repetitive and many deaths occur for reasons other than competitors' innovation.

3.3.3 Advantages of Establishment Birth Measures

The advantages of dynamic data are that they are flow data; they better capture innovation and dynamic micro data can be refined to include only the most innovative establishments. Flows measure the change over a particular period of time and is less related to the stock of establishments, which is not a particularly good proxy for entrepreneurship, because the stock is taken at one point in time and gives us no information about innovation, success, or longevity. Another good use of establishment birth data is to refine the data to include the most innovative firms (Luger and Koo, 2005; Mueller, 2007).

3.3.4 Disadvantages of Entrepreneurship Birth Measures

Disadvantages of using dynamic data include finding an appropriate denominator and period of time, and accessing dynamic data. Dynamic data are more costly and time-consuming to use than stock data, which are generally publically available. The Census' Longitudinal Business Database (LBD) contains the universe of firms and allows estimation of births and deaths, however, accessing these data is a lengthy and costly process that researchers can pursue, but practitioners are not able to gain access to or use for regional economic development benchmarking and policymaking.

The period of time in which flows are examined can affect results, especially across different points of the macroeconomic cycle, making the measure extremely volatile year-to-year (Spelman, 2006; Tamasy, 2006). Lee et al. (2004) find the 2000 birth rate is 1.16 to 5.05 per 1000 people in U.S. Metropolitan Statistical Areas but Acs and Mueller (2008) find the rate ranges from three to 18 over 1998-2001, illustrating the volatility. Finally, there is disagreement as to the appropriate denominator for dynamic data. Regional studies use both population and employment with slightly different results, particularly among heterogeneous units of observation. Macroeconomic studies usually use establishments.

3.4 OTHER PROXIES FOR ENTREPRENEURSHIP

3.4.1 Income

Proprietor income is a measure of the economic value of the self employed to an economy and serves as a proxy for entrepreneurial success (Low, 2004; Goetz and Shrestha, 2009). The user assumes that as average proprietor income rises, the region as a whole becomes more prosperous. The self employment income data are problematic, however, because, although based on IRS tax filings, the BEA-REIS data are highly imputed. The BEA adjusts income up by as much as 40 percent to account for under-reporting of income. This and other adjustment procedures, some of which are not specified by BEA, make the data suspect. Finally, using income data without accounting for cost-of-living is problematic, making self employment income a poor choice to measure the value of entrepreneurship, particularly because a lot of self employment is by necessity (Reynolds et al., 1994).

3.4.2 Patents

Measures of invention, while certainly possessing the "creative" portion of Schumpeter's entrepreneur, are not measures of entrepreneurship. Patents fail as a measure of entrepreneurship because there is no firm and we do not know if the inventions make it to market; despite this, entrepreneurship literature routinely uses patent data (Wong et al., 2005, Trajtenberg et al. 2006). In addition, research facilities, universities, and high-tech firms, which are most likely to generate patents, likely exist in populated areas, thereby creating an endogeneity problem for researchers interested in teasing out the causality between innovation, place, and entrepreneurship (OhUallachain, 1999; Carlino et al., 2007).

3.4.3 Trait Approach

Individuals' traits, identified in surveys, have been used to measure entrepreneurship (Bull and Willard, 1993). Low and MacMillan (1988) conclude that there is no typical entrepreneur and that attempts to profile such a person are futile because entrepreneurs are, by definition, atypical people. Following Low and MacMillan, researchers are moving away from trait-based measures to functional measures of entrepreneurship.

3.5 EMPIRICAL COMPARISON OF MEASURES

Table 3.1 shows entrepreneurship attributes of measures discussed in this chapter, their relationship to my three attributes of entrepreneurship, summary statistics, source, and definition. Measures in Table 3.1 represent the most widely-used regional entrepreneurship measures and their analysis and comparison in this chapter is summarized here; none of the measures are ideal, rather, their use appears to be based upon their availability.

The self employment measure ignores innovation, but has a positive relationship with growth. The mean of the nonfarm self employment rate is 0.25—that is one quarter of nonfarm employment in U.S. counties is in self employment, with no employees—a relatively high rate which is exacerbated by the inclusion of multiple job holders and part-time proprietorships. As expected with change in a stock measure, the change in the nonfarm self employment rate is very small, however, the county-to-county distribution of the nonfarm self employment rate is large.

The establishment measures ignore innovation and risk/uncertainty bearing, making them the least entrepreneurial and, likely due to this, establishment rates do not have a statistically significant relationship with employment growth (Table 3.1). Dynamic establishment data are the closest to my posited definition of entrepreneurship and has a positive relationship with growth. The measure varies a lot over counties though; aggregate establishment births, for all counties, have a larger range than in metropolitan statistical areas, as reported in Lee et al., 2004 and Acs and Mueller, 2008.

The omission of innovation in most of the measures is striking (Table 3.1). Summary statistics and relationships with growth suggest that the most promising measures of entrepreneurship, currently available, are the self employment rate and dynamic establishment data.

Table 3.1 Comparison of Entrepreneurship Measures

	Owner/Operator		Risk Bearer		Innovator						
Entrepreneurship measure description	ownership	operation	risk	uncertainty	innovation	reallocation	Mean	StDev	Min	Max	Corr with Growth^
Nonfarm proprietors over nonfarm total employment*	X	X	X	X			0.247	0.094	0.030	0.710	+
Change in proprietor rate, 2001-2006*	X	X	X	X			0.010	0.005	0.000	0.071	+
Establishments over population**/*		X					0.024	0.009	0.004	0.116	
Employees over estabs**		X					0.332	0.100	0.015	1.100	
Single-unit establishment births over emp (1000)***	X		X	X	X		9.401	6.028	0.000	96.774	+
Single-unit establishment deaths, over emp (1000)***	X		X	X		X	0.073	0.020	0.000	0.500	+
Average nonfarm proprietor income*	X	X					0.125	0.065	-0.725	0.462	-
Patents over population****/*				X	X		0.0001	0.0003	0.0000	0.0033	+

*Bureau of Economic Analysis, Regional Economic Information System, 2006
**County Business Patterns, US Census Bureau, 2006
***US Census Bureau, Statistics of Businesses, 2002-2003
****United States Patent and Trademark Office, 2006
^Employment Growth, 1991-2006, significant at 0.001 level

3.6 CONCLUSION

I presented different entrepreneurship measures, their relationship to my definition of entrepreneurship, and advantages and disadvantages of their use in this chapter. I also discussed their variation across space and relationship with economic growth. I find self employment is the most widely used measure of regional entrepreneurship, due to its availability over time and space. Self employment, however, is not an ideal measure of regional entrepreneurship because it grossly overestimates entrepreneurship by ignoring innovation and several important measurement issues exist with U.S. county-level data, which must be carefully considered.

Establishment rates ignore both innovation and risk/uncertainty bearing, making them less of an entrepreneurship measure than a proxy for past entrepreneurship, however, it is a useful measure of entrepreneurship over long periods, and is widely available for use. Use of dynamic establishment data is growing as more micro- and dynamic datasets become available to researchers. Establishment birth data capture the firm ownership and risk bearing attributes of entrepreneurship and the innovation attribute, to a certain extent, because they are flow data. More importantly, dynamic data are not publically available, thus, unusable by practitioners and policymakers.

The current state-of-the-art in measuring regional entrepreneurship is a hodge-podge of second-best measures, based upon available data, and with no consensus among researchers, economic development practitioners, or policymakers. Most troubling, the commonly used measures of entrepreneurship ignore innovation—a long established defining attribute of entrepreneurship that drives economic development (Schumpeter and Opie, 1983). No one measure discussed in this section meets the three attributes of my entrepreneurship definition. Indeed, one measure cannot be expected to measure individuals and firms, stock and flow, change, ownership, risk-bearing, and innovation (Gartner and Shane, 1995). Indices have been used to combine one or more dimensions of entrepreneurship into one measure (Iversen et al., 2008) but indices of entrepreneurship are fraught with weighting and measurement problems of their own (Eff, 2007). In the next chapter, I develop a method for identifying innovative components of self employment and establishment births, the more promising of the measures discussed in this section.

CHAPTER 4: THE ENTREPRENEURIAL INDUSTRIES INDICATOR

This chapter responds to a call in the entrepreneurship literature for the development and dissemination of reliable entrepreneurship metrics (Baumol, 1993; Gartner and Shane, 1995; Goetz and Freshwater, 2001; Glaeser, 2007). Better indicators can improve entrepreneurship research, add value to practitioners' economic development work, and make entrepreneurship policies more effective.

I develop a new indicator of entrepreneurship that captures the innovative dimension of entrepreneurship ignored by others. Identifying innovative industries is the key contribution. Combining indicators of innovative industries with federal statistics on self employment and establishment births creates an indicator of entrepreneurship for all counties, *Entrepreneurial Industries*. This is the first indicator to capture all the dimensions of entrepreneurship (Figure 4.1)

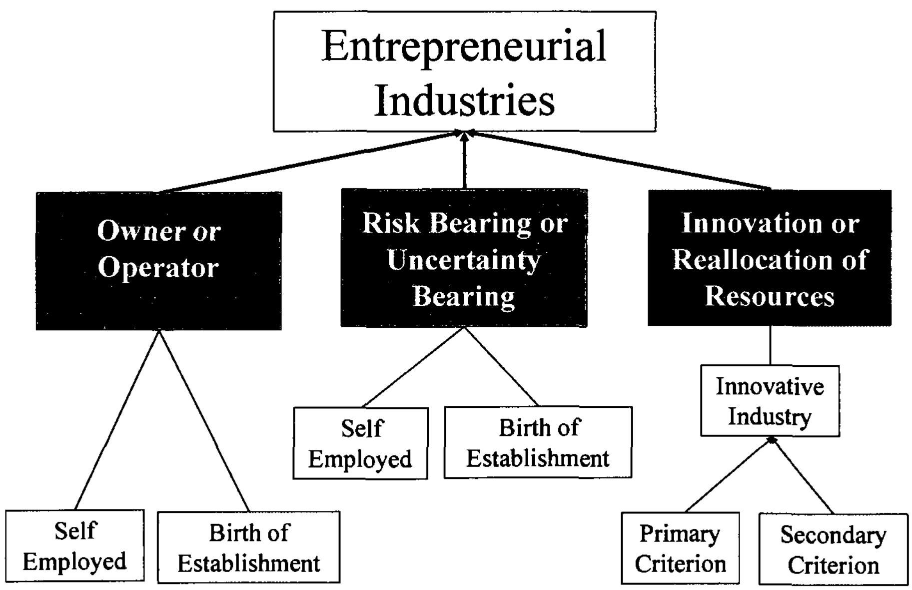

Figure 4.1 Entrepreneurial Industries and Their Specification

I define innovative industries as meeting one primary and one secondary criterion of innovation. The two primary criteria are technology and skill. I use high technology as a primary indicator of innovative industries because high tech industries are considered more innovative and more apt to use emerging technologies than other industries. I use high skill as a primary indicator of innovative industries because research has identified a link between skills and innovation (Yemen and Lahr, 2008). The five secondary criteria are a lower threshold of high skill or technology, patents, churn, and employment growth, but innovative industries must only meet one secondary criterion.

I create an indicator of entrepreneurship that incorporates innovation using the resulting innovative industries. I count the number of innovative industry establishments in both self employment data and a special tabulation of single-unit (non-branch, independent) employer establishment births data. Both are available annually at detailed industry levels. I standardize the resulting count to obtain *Entrepreneurial Industries*, which is the first annual, county-level indicator of multiple facets of entrepreneurship, including innovation.

In this chapter, I show that Entrepreneurial Industries is a conceptual and empirical improvement over other entrepreneurship indicators and measures. The nexus between innovative industries and self employment and establishment births makes Entrepreneurial Industries useful. I also show that Entrepreneurial Industries is robust to changes in the innovative industry definition, suggesting that the method is effective even if the inclusion of certain industries might be surprising.

I proceed by discussing the criteria and method used to identify innovative industries and create the Entrepreneurial Industries indicator. In the results section, I describe Entrepreneurial Industries and support its empirical validity by comparing it to widely used entrepreneurship indicators that are not available for most counties. Finally, I demonstrate robustness and discuss the merits of using it as a regional indicator of entrepreneurship.

4.1 CRITERIA FOR IDENTIFYING INNOVATIVE INDUSTRIES

4.1.1 High Technology Industries

Many definitions of high tech exist for both occupations and industries. They vary widely and are difficult to quantify. For instance, the U.S. Census Bureau defines high tech occupations as those *embodying new or leading edge technologies*. The Congressional Office of Technology Assessment describes high tech industries as those *engaged in design, development, and introduction of new products and/or innovative manufacturing processes through the systematic application of scientific and technical knowledge*. Others use judgment to identify high technology industries (Niosi, 2000). Defining high tech may be as difficult as defining entrepreneurship.

I adopt the Bureau of Labor Statistics' (BLS) empirical definition, which uses percent of industry employment in high tech occupations. Occupations, not industries, is the base unit because many workers in high tech industries do not utilize technology in their work, e.g., administrative assistants or marketing specialists; including such workers overstates the extent of high tech activity in these industries (Kilcoyne, 2001). Defining high tech industries with average education also proves problematic, e.g., percent of employees who hold a college degree in science or engineering (Mueller, 2008). This method tends to identify high wage occupations rather than high tech because it includes occupations which utilize technology that has been available for generations, e.g., process engineers, while omitting jobs directly related to the concept of high technology, e.g., technicians (Kilcoyne, 2001).

The Bureau of Labor Statistics' (BLS) definition of high tech occupations includes science, engineering, or technology-oriented technicians and workers who typically use new technologies to perform their duties (Table 4.1), to identify high tech industries. The BLS also provides the Standard Occupation Classification (SOC) codes for high tech occupations, enabling me to calculate employment by occupation by industry.

Table 4.1 BLS High Tech Occupations, 2000

Computer and information scientists, research	Nuclear engineers
Computer software engineers, applications	Petroleum engineers
Computer software engineers, systems software	Aerospace engineers
Geological and petroleum technicians	Biomedical engineers
Network systems and data communications analysts	Chemical engineers
Electronics engineers, except computer	Electrical engineers
Mining and geological engineers	Chemists
Aerospace engineering and operations technicians	Astronomers
Electrical and electronic engineering technicians	Physicists
Electro-mechanical technicians	Microbiologists
Geoscientists, except hydrologists and geographer	Biological technicians
Multi-media artists and animators	Chemical technicians
Medical and clinical laboratory technologists	Computer systems analysts
Nuclear medicine technologists	Nuclear technicians
Radiologic technologists and technicians	Epidemiologists
Medical scientists, except epidemiologists	Database administrators
Atmospheric and space scientists	Computer programmers
Computer hardware engineers	Biochemists and biophysicists

Source: BLS, 2001

To define high tech industries I calculate percent employment in high tech occupations (Bednarzik, 2000). The Industry-Occupation National Employment Matrix, (Employment Matrix) contains employment in each occupation for each industry, enabling me to calculate percent employment in high tech occupations for each industry. The Employment Matrix contains data for each SOC occupation for industries at the four- to six-digit NAICS industry level. I use the 2006 Employment Matrix, which includes over 300 industries (2002 NAICS) and over 700 SOC occupations. Industry-occupation employment cells that are confidential, having fewer than 50 jobs, or are of poor statistical quality, are suppressed.

I calculate the average percent high technology employment is 3.5 percent, standard deviation, 7.1 percent. Industries at the high end of the distribution were in the Information (NAICS 51) and Professional, Scientific, and Technical Services (NAICS 54) sectors. For example, "Software Publishers" (NAICS 51121) had 42 percent high tech employment, "R & D in the physical, engineering, and life sciences" (NAICS 54171) had 35 percent high tech employment, and "Testing Laboratories" (NAICS 54138) had 27 percent high tech employment.

To differentiate "high tech" industries from others, BLS uses a cutoff scheme based upon the mean of high tech employment, defining industries with at least three times the mean level (10.5 percent) of high tech employment as "medium-content" high technology industries (Bednarzik, 2000; Hecker, 2005). Only 19 industries, 3.8 percent of 5-digit industries, meet this criterion, and they are primarily in Manufacturing, Information, and Professional, Scientific and Technological Services sectors. I use the BLS definition of high tech industries because it is a very strict criterion, which ensures high tech industries differ from all industries and differentiates Entrepreneurial Industries from its parent measures and other metrics.

BLS defines industries with two times the mean level (7 percent) of high tech employment as "low-content" high technology industries. Thirty industries, or six percent, meet this, lower, criterion; the eleven additional industries are primarily in the Manufacturing sector and Transportation sector, which suggests that by lowering the cutoff, less technological industries are included. I use two times the mean of high tech in subsequent sensitivity analysis of the Entrepreneurial Industries method.

One caveat with this method is that high *tech* industries do not necessarily imply *innovative* industries. By definition, innovation is the creation of a new product or process and high tech industries, by definition, are engaged in design, development, and introduction of new products and/or innovative manufacturing processes. Given the similarities between these definitions, I argue that high tech industries are a reasonable proxy for innovative industries and, by necessitating a secondary innovation criterion, I capture only the most innovative industries.

4.1.2 Identifying High Skill Industries

As an alternative to high tech, I use high skill as a primary indicator of innovative industries because research has linked innovative and entrepreneurial activities to high skill employees (Lee et al., 2004; Mueller, 2007; Munn, 2008). A higher skilled workforce has the necessary tools to create new products and processes. In this section, I describe how I identify high skill occupations using ONET data. I choose occupations that have the highest level of skills and knowledge that generate innovation and product creation, e.g., problem solving, critical thinking, science and engineering knowledge. Using my high skill occupations, I identify high skill industries using the same method used to identify high tech industries using high tech occupations.

4.1.2.1 Identifying high skill occupations

The advent of occupation-based data and categorical schemes to organize occupation-level data has increased research using high skill occupations (Feser, 2003; Koo, 2005; Yemen and Lahr, 2008). Identifying high skill occupations is preferable to using proxies like education, e.g., Mueller (2007), because unlike education levels, occupation provides more information about the actual duties of an employee.

Occupation-based data are available in ONET-SOC, the Occupational Information Network survey that uses Standard Occupation Classifications (SOC) titles to match occupations to their attributes, such as skills. The ONET-SOC 12.0 (2006) database contains survey data on occupational attributes for 949 SOC occupations and includes comprehensive information on worker attributes, including skills, knowledge, and education for each occupation. I use the Worker Requirement module of the ONET-SOC to assess information on "Skills" and "Knowledge" for each occupation. The other modules are Worker Characteristics, Workforce Characteristics, Occupational Requirements, Experience Requirements, and Occupation-Specific.

I select 20 skill and knowledge categories that are relevant to identifying employees in innovative industries, using BLS' definition, i.e., engineering or technical skills or skills that include qualities essential to the process of innovation (Table 4.2). I include knowledge categories in addition to skill categories because relatively few occupations had high levels of skills since they are more general than knowledge. I found that including the knowledge categories increased the scientific and technological skills of the occupation set, essential for capturing high skill occupations principally involved in innovation.

Table 4.2 Selected ONET Skill and Knowledge From Worker Requirement Module

Skill	Knowledge
Critical Thinking	Computer and Electronics
Time Management	Mathematics
Complex Problem Solving	Telecommunications
Programming	Engineering and Technology
Technology	Communications and Media
Science	Chemistry
Writing	Design
Speaking	Physics
Operations Analysis	Biology
Troubleshooting	

Data for each occupation on both the level of and importance of each skill/knowledge are available from ONET. For each occupation, the skill/knowledge levels, **V**, are on a scale of 0 to 7 and the importance of the skill/knowledge to the occupation, **M**, is on a scale of 1 to 5.

To integrate the level of skill, **V**, and its importance, **M**, into one metric, I create a weighted matrix using Feser's (2003) method. The level of skill/knowledge for each occupation is V_{ij}, where the ONET survey data give a level, V, of skill/knowledge, *j*, for each occupation, *i*, and importance, M, of each skill/knowledge to each occupation is M_{ij}. Weighted matrix $\boldsymbol{S}$ relates the importance, **M**, to the level, **V**, of each skill; let $\boldsymbol{S}=\boldsymbol{V}*\boldsymbol{M}$, where $S_{ij} = V_{ij} * M_{ij}$. By taking the product of V_{ij} and M_{ij}, Feser most heavily weights knowledge that is of both a high level as well as central to the occupation. For example, for *i*=Economist and *j*=Critical Thinking, V=5.48 (out of 7) and M=4.56 (out of 5)—both are relatively high—but for *j*=Chemistry, V=0.8 and M=1.3, illustrating that chemistry knowledge is unnecessary and unimportant. Thus, for economists, where *j*=Critical Thinking, S_{ij}=5.46*4.56=24.99 (out of 35) but where *j*=Chemistry, S_{ij}=0.8*1.3=1.04 (out of 35).

I classify an occupation as *high skill* if its S_{ij} value is high enough to meet the cutoff for any one of the 20 selected skills, resulting in 119 high skill occupations. I define the cutoff value for each of the 20 selected skills, j, based upon the distribution of the S_j values. I define "high" for each skill as a S_i value greater than three standard deviations above the mean. Where the tails of the distribution are small—so small that no S_i values were greater than three standard deviations above mean—I use two standard deviations for the cutoff (Complex Problem Solving, Critical Thinking, Speaking, Time Management, and Writing). I use standard deviation rather than a multiple of the mean because most of the skills follow a Normal distribution and I wanted to make the criteria difficult to meet, so three standard deviations above the mean includes only 0.3 percent of occupations for each skill. For example, Critical Thinking has a cutoff value of 28.48, recalling that the S_{ij} where i=Economist and j=Critical Thinking is 24.99, we know economist does not meet the high skill criterion via Critical Thinking. The number of occupations classified as *high* in each of the 20 skill/knowledge fields ranged from four (Operations Analysis) to 64 (Biology). Of the 119 occupations classified as high skill, many of those occupations met the *high* threshold for several skills.

4.1.2.2 Identifying high skill industries

Having identified high skill occupations, I use the Employment Matrix to calculate percent high skill employment for each industry, just as I did for high tech (Figure 4.2). Again, I use three times the mean level of percent high skill employment as a cutoff to define high skill industries because the high cutoff value leaves the most skilled industries, which is necessary to differentiate innovative industries from all industries and Entrepreneurial Industries from its parent measures. Three times the mean, 17.2 percent, is the cutoff to designate a *high skill industry*; percent high skill employment for each industry ranged from 0.004 percent to 58 percent

Figure 4.2 Process for Selecting High Skill Industries

One caveat is that the employment matrix does not include self employment by occupation and industry. Consequently, I must assume that high skill industries, defined by paid employees, are high skill industries for the self employed.

Assumptions used to identify high tech occupations are also worthy to note. Using 20 skill/knowledge attributes, and requiring occupations to meet only one, adds breadth to the high skill occupation definition and results in almost one in nine occupations being high skill. I use a broad definition to define high skill occupations because I restrict the number of high skill industries by making that cutoff high, three times the mean level of high skill employment. Like high tech, I set the cutoff very high so that only the highest skill industries are included. This cutoff enables me to differentiate high skill industries from the universe of industries.

My method for identifying high skill industries is similar to the high tech industry method. I believe the high skill and high tech indicators of innovation are superior to the secondary criteria discussed in the next section, but, by necessitating a secondary innovation criterion, I hope to capture only the most innovative industries, which is necessary to differential Entrepreneurial Industries from other entrepreneurship indicators.

4.1.3 Identifying High Patenting Industries

Patents have been widely used to measure invention (Wong et al., 2005; Trajtenberg et al., 2006), and, interacted with establishments, can measure Schumpeterian innovation, *creating a new product, process, or service within an organization.* Both traits make patents a useful secondary indicator of innovation. Because patent data have important flaws, I do not use patents as a primary criterion. The combination of high tech or high skill and patenting suggests invention and innovation occurs simultaneously (Munn, 2008).

Data are available from the United States Patent and Trade Office (USPTO). Patents granted in a single year by county are relatively random due to the scarcity of patenting. Consequently, I use patents granted between 1990 and 1999, the most recent available data. In addition, I only use patents assigned to non-government organizations and individuals (U.S. or Foreign) because I use private-industry data throughout the dissertation.

There are 417 patent classes, a relatively high level of detail. Unfortunately, patent classes cannot be translated directly into NAICS industry sectors. The only link between patents and industries are via 1972 Standard Industrial Classification (SIC) codes. The USPTO provides patent classes and product field titles from the Office of Technology Assessment and Forecast (OTAF), which creates a link between OTAF fields and the 1972 SIC.

To get a NAICS code for high patent industries, I match patent classes/OTAF fields to 1972 SIC codes and then use a SIC-NAICS bridge to assign NAICS industry codes to patent classes. The method for matching SIC to patent class is broad and outdated. Technological developments between the 1972 SIC and the 2005 patent classification scheme leave wide gaps in industries that have grown substantially over the past 40 years, e.g., Typewriters and Office Computing. More importantly, the match between the SIC and OTAF proves imperfect because the OTAF tables do not consider all relevant patents or exclude all irrelevant patents. To improve the match, I use patent class descriptions and NAICS codes descriptions to clarify which NAICS code should be used. Silverman (1999) and Porter and Stern (2003) used an algorithm to match patent classes to 6-digit NAICS code, but their algorithms are not publicly available.

I identify high patent industries by summing the number of patents granted in each patent class and selecting a cutoff value to define high patent classes. Patents are count data and have a Poisson distribution. Given the shape of the distribution, I define high patent classes at the natural break in the distribution of the data—the tail of the Poisson distribution–because the mean is meaningless. Patent classes had considerable overlap across NAICS, leaving 32 high patent industries, of which 24 are manufacturing. Industry codes for some patent classes in the tail could not be identified, and as a result, are omitted from the high patent industry list. Omitted industries were more likely to be newer industries. I am not worried about this small bias because the patent criterion is only secondary to high skill or high tech.

Again, patent data have many problems, as many inventions are not patented. The degree of incremental patenting varies by industry, and the economic impact of patenting varies for regions and industries (Carlino et al., 2007). Patent data are likely to overestimate invention in industries that incrementally patent and underestimate innovation in others. The manufacturing sector is more likely to patent than other sectors that may be equally innovative (Orlando and Verba, 2005); 75 percent of high patent industries are manufacturing industries even though manufacturing represents only 4.2 percent of all establishments. Finally, even when normalized by regional population, patent rates are correlated with population and systematically less likely in rural regions (OhUallachain, 1999).

4.1.4 Identifying High Churn Industries

The churn rate has been widely regarded as a key measure of Schumpeter's *creative destruction,* making it an appropriate indicator of innovativeness. Defined as the sum of the establishment birth rate and the death rate, the churn rate captures the continual reinvention of products, practices, and services (Peneder, 2008; Iversen et al., 2008). The birth rate is an indicator of innovative or cost-effective ideas and the death rate is an indicator of the firms made obsolete by births or acquisitions.

Data used to calculate the churn rate are from a special tabulation of the Statistics for U.S. Businesses, Bureau of the Census, courtesy of USDA, Economic Research Service; these are the same data I use for establishment births (see Appendix A). I use 2000-2003 to calculate the churn rate because, much like patents, the births occurring in one year are relatively random. I use births and deaths for single-unit establishments because these establishments are considered more entrepreneurial than branch units, whose entry/exit is decided by a distant Headquarters facility.

I sum the single-unit employer establishment birth rate and death rate, nationally, for each five-digit NAICS code. I calculate the birth rate and death rate for each industry using the total number of establishments in each NAICS as the denominator (Equation 4.1), which is consistent with the way others have calculated the churn rate (Peneder, 2008; Iversen et al., 2008). The mean churn rate is 0.15 and the median is 0.28, and I use the median, approximately twice the mean, as a cutoff to define the secondary criterion because three times the mean, resulted in very few “high churn” industries, while using the mean resulted in the majority of industries being high churn. Although using the median makes half of all industries high churn, I think this is an appropriate cutoff for a secondary criterion.

$$(4.1) \qquad Churn = \frac{\sum_j births_{ij}}{\sum_j estabs_{ij}} + \frac{\sum_j deaths_{ij}}{\sum_j estabs_{ij}}$$

Whereas some industries have regulatory or institutional barriers to high churn, e.g., Finance and Insurance (NAICS 52), industries with lower barriers to entry often have higher churn rates. Professional, Scientific and Technical Services (NAICS 54), has a 0.6 churn rate—four times the average and over twice the median. Average employment in the industry is three, suggesting that these establishments are staffed by a small number of professionals and/or support staff, some of whom could be part-time employees.

4.1.5 Identifying Industries with the Innovation Stage of the Profit Cycle

The empirical definition of the *innovation stage* of the Profit Cycle is based upon Ann Markusen's (1985) *Profit Cycles, Oligopoly, and Regional Development*. The Profit Cycle model organizes information about the timing of industries' lifecycle stages—one of which is innovation. Profit Cycle is closely related to the Product Cycle (McDonald and McMillen, 2006). I proceed with a brief discussion of the Profit Cycle, and then discuss the data and methods in its use as a secondary criterion.

The Product Cycle answers Vernon's (1966) call to interpret the timing of innovation and the decentralization of production (Norton and Rees, 1979). It relies on the notion that industries have defined lifecycle stages. The five stages of the profit cycle include *zero profit or experimentation* (initial firm birth, product design), *innovation or super-profit* (profits/rents from an innovative edge), *mature or normal profit* (market saturation), *concentration* (competition or oligopoly), and *negative profit/death*. The *innovation* stage captures product innovation, when firms make product improvements, perfect production, and drive down the cost of production through innovation. Additionally, during this stage, the lack of competition allows for high prices while the industry is growing (Markusen, 1985). The innovation stage of the profit cycle captures this concept, entrepreneurs gaining a monopoly position. Since the *innovation* stage identifies when an industry has the highest profits from an innovative edge, I assume that being in the *innovation* stage of the profit cycle is a useful secondary criterion for identifying innovative industries. Process innovation, however, tends to occur in the *mature* and *concentration* stages of the cycle and because I focus on product innovation, is not included in this criterion.

Markusen defines firms in the innovation (super-profit) stage as having an average annual employment growth rate, between observations of a smoothed series, greater than two percent (>2%). The lengths of profit cycle stages vary. Sorenson (1997) identified profit cycle stages that span multiple decades, but these vary by industry. Consequently, I focus on employment growth during the 2001 to 2007 peak-to-peak business cycle (March 2001-December 2007) in order to capture industries currently in the innovation stage, and that maintained or gained that status during that period. This business cycle is particularly useful because of the unusual growth the economy exhibited during the 1991-2001 business cycle, which put many industries in Markusen's high growth stage (greater than two percent average annual employment growth).

I use employment data from the NAICS based Quarterly Census of Employment and Wages (QCEW). The QCEW data are based on unemployment insurance records, and include data on paid employees, but not the self employed. The six-digit NAICS data are national employment totals for each month, 1990 to 2007. I smooth average annual employment data to purge the data of random effects and national business cycle effects using method developed by Neumann and Topel (1991) and Sorenson (1997). I smooth and plot the industries with more than two percent average annual employment growth and no irregularities. I remove industries with irregular plots or other data-induced abnormalities from the analysis.

I find 29 percent of five-digit NAICS industries meet Markusen's criteria, including many service industries (37 percent of all industries in the innovative stage of the cycle) and manufacturing industries (45 percent). Although Markusen's definition identifies innovative industries, and is grounded in economic theory, the measure includes industries that are growing for non-innovation reasons, such as consumer preference. Requiring innovative industries meet both a primary criterion and a secondary criterion reduces overestimation of innovative industries.

4.1.6 Rejected Criteria for Identifying Innovative Industries

I considered using industries newly recognized in NAICS or national input-output tables as a secondary criterion, but many are not innovative. While some industries are recognized for the first time in NAICS because of innovation, e.g., satellite communications and software reproduction, many non-innovative industries are new, e.g., bed and breakfast inns, pet supply stores, and diet centers. These industries capture today's changes in preferences, rather than innovation. There is no way to distinguish the new and innovative from the new and non-innovative industries, for instance, fiber optic cable manufacturing, limited service restaurants, and convenience stores are new industries in NAICS, but deciding which are innovative is difficult. A similar argument can be made for not using the 1997 and 2002 input-output codes. The majority of changes were in declining industries rather than innovative industries.

Gazelle establishments, those exhibiting rapid growth in employment and revenue growth, are not used to identify innovative industries because such establishments exist in *all industries* (Acs, Parsons, and Tracy, 2008). More importantly, identifying Gazelle establishments at the five-digit NAICS is difficult because data on revenue for detailed industry sectors are not available (Birch, 1981). Although employment growth data for industries are available, using these data would replicate the innovation stage criterion.

4.2 METHOD

In this section, I discuss the method used to identify innovative industries and the method I use to create Entrepreneurial Industries, ST3, named because a key component of the method is that Skill or Tech employment must be three times the mean.

4.2.1 Identifying Innovative Industries

I identify innovative industries at the five-digit NAICS industry level. Construct validity decreases with aggregation, but the Employment Matrix is not available at the six-digit level for many industries, making five-digit NAICS is the lowest usable level of aggregation. Using four-digit NAICS resulted in many overlapping industries, an unsuitable level of aggregation. To meet the primary criterion, an industry must have at least three times the average level of percent high skill employment (17.2 percent), or three times the average level of percent high tech employment (10.4 percent). Therefore, I call this method ST3 (Skill or Tech at three times the mean). Both are very selective cutoff levels by design; I chose these levels, because BLS uses these levels and using a high cutoff ensures that relatively few industries are either high tech or high skill, distinguishing my indicator of entrepreneurship differs from others. In Section 4.4., I examine the sensitivity of this choice by testing two alternative methodologies for identifying innovative industries, including lower the criterion to two times the mean for both skill and tech.

I use more relaxed standards for the secondary criterion. The secondary criterion acts only as innovation insurance after passing the primary hurdle. An industry can meet the secondary criterion using Patent, Churn, or Profit Cycle criteria, as discussed in 4.1, or, by exceeding the average level of skill or tech employment (although if an industry's primary criterion is high skill, it's secondary criterion cannot be skill and vice-versa). I use the average for skill and technology as a secondary criterion because it represents a relatively high standard that is below the cutoff for "high," but is well above the median. Using two times the average was also a difficult standard to meet, so I use the average because only one-quarter of industries exceed the average percent high skill employment (5.7 percent) and the average is almost three times the median (2.0 percent). Similarly, the average percent high technology employment is 3.5 percent and the median is 0.6 percent.

Summarizing, the five secondary criteria are: the patent criteria described in Section 4.1 (industry is in the top 15 percent of patent activity), the churn criteria described in Section 4.1 (the churn rate, birth rate plus death rate, is greater than its median or two times the average), the innovation stage of the Profit Cycle described in Section 4.1 (average annual employment growth over the 2001-2007 business cycle is greater than 2 percent), the percent high skill employment of at least its average (almost three times its median), and the percent high tech employment of at least its average (almost six times its median).

The combination of primary and secondary criteria allows for the identification of the *most* innovative of industries. I argue that using multiple identifiers of innovation leaves only the most innovative industries, differentiating Entrepreneurial Industries from other entrepreneurship measures. This multi-criteria strategy is not original; Peneder (2008) does the same, arguing that a combination of two identifiers better captures truly entrepreneurial establishments.

4.2.2 Resulting Innovative Industries

The ST3 method identified 39 innovative industries (Table 4.3). High tech was the primary indicator for 19 industries and high skill for 18 industries. No industries had high levels of both skill and tech, likely because a high level of specialization in one or the other is necessary to meet the cutoff. I believe this indicates the demanding nature of the criteria—that innovative industries are different from the universe of industries.

Many of the innovative industries met two or more secondary criteria, e.g., Medical and Diagnostic Laboratories, NAICS 62151. This is likely because the standards for secondary criteria are much lower than for the primary criterion, which allows them to be met more easily. If cutoffs for the secondary criteria were as high as for the primary criteria, very few industries would qualify as innovative industries.

Innovative industries are primarily in Manufacturing (NAICS 31-33), Information (51), and Professional, Scientific, and Technological Services (54) (Table 4.3). These sectors are overrepresented compared to their share of total industries. Twelve, or 32 percent, of innovative industries are manufacturing industries but manufacturing establishments comprise only 4.2 percent of private establishments (Q1:2008, QCEW). The Information sector represents 16.2 percent of innovative industries, but only 1.7 percent of private establishments, and 16.0 percent of the innovative industries are in the Professional, Scientific, and Technological Services sector but only 11.6 percent of establishments are in this sector.

I exclude Mining (NAICS 21) industries from this analysis due to their year-to-year volatility and dominance in particular regions. The only industry in this sector that would have been an innovative industry is Oil and Gas Extraction (NAICS 2111). Peaks and valleys in oil and natural gas prices cause lots of entry and exit, which affects annual birth data for employers. This sector also has dramatic volatility in year-to-year self employment in regions with many independent oil pumps.[4]

[4] I consulted with Mike Orlando, a regional economist with expertise in the energy industry, and former Shell Oil engineer. Orlando informed me that several regions in the country are rich in *independent oil and gas producers,* sole proprietors who "turn on the pump" when oil or gas prices rise to a certain level. This practice creates year-to-year volatility in births/deaths and self employment in the Oil and Gas Extraction sector, particularly in West Virginia, Oklahoma, and the South.

Table 4.3 Innovative Industries and Composite Criteria

Naics	Description	High Skill	High Tech	Ave Skill	Ave Tech	Churn	Patent	Profit Cycle
22110	Electric Power Generation, Transmission and Distribution	X		X	X		X	X
31161	Animal Slaughtering and Processing	X		X		X		X
32510	Basic Chemical Manufacturing		X		X	X	X	X
32541	Pharmaceutical and Medicine Manufacturing		X		X	X	X	X
33411	Computer and Peripheral Equipment Manufacturing		X		X	X	X	
33421	Communications Equipment Manufacturing		X		X	X	X	
33431	Audio and Video Equipment Manufacturing		X		X	X	X	
33441	Semiconductor and Other Electronic Component Manufacturing		X		X	X	X	
33451	Navigational, Measuring, Electromedical, and Control Instruments Manu.		X		X	X	X	
33461	Manufacturing and Reproducing Magnetic and Optical Media		X		X	X	X	
33641	Aerospace Product and Parts Manufacturing		X	X	X	X	X	X
33661	Ship and Boat Building	X		X				X
33711	Wood Kitchen Cabinet and Countertop Manufacturing	X		X				X
42370	Hardware, Plumbing and Heating and Supplies Merchant Wholesalers	X		X		X		
44110	Automobile Dealers	X		X		X		
51121	Software Publishers		X		X	X		
51611	Internet Publishing and Broadcasting		X		X	X		
51711	Wired Telecommunications Carriers		X		X	X		
51731	Telecommunications Resellers		X		X	X		
51811	Internet Service Providers and Web Search Portals		X		X	X		
51821	Data Processing, Hosting, and Related Services		X		X	X		
54121	Accounting, Tax Preparation, Bookkeeping, and Payroll Services	X		X		X		
54138	Testing Laboratories		X		X	X		
54151	Computer Systems Design and Related Services		X		X	X	X	
54171	Research and Development in the Physical, Engineering, and Life Sciences		X		X	X		
54172	Research and Development in the Social Sciences and Humanities	X		X	X	X		
54194	Veterinary Services	X		X		X		X
55111	Management of Companies and Enterprises	X		X	X	X		
56111	Office Administrative Services	X		X	X	X		X
56142	Telephone Call Centers	X		X		X		
56151	Travel Agencies	X		X		X		
56161	Investigation, Guard, and Armored Car Services	X		X		X		
56190	Other Support Services	X		X		X		
62151	Medical and Diagnostic Laboratories		X	X	X	X		
62211	General Medical and Surgical Hospitals	X		X	X		X	
62420	Community Food and Housing, and Emergency and Other Relief Services	X		X		X		
81121	Electronic and Precision Equipment Repair and Maintenance	X		X		X		X

4.2.3 Caveats

The industries in Table 4.3 *represent* innovative industries, but some industries surely could be included and others excluded. Innovative industries should be interpreted as part of a method to develop a better indicator of entrepreneurship—the cost of developing such a method is that some industries are included and some are not. The stringent criteria is what makes innovative industries different from all industries, and, when applied to self employment and establishment birth data, will result in an indicator of entrepreneurship that is truer to my conceptual definition of entrepreneurship.

Innovative industries should be interpreted as a whole and not be used for targeted recruitment because it is not an exhaustive list, rather an indicator of innovation at the industry level. Innovative industries should be interpreted similarly to Richard Florida's Creative Class; Florida includes occupations in which creative people are *most likely* to work, it is not a finite list of who is creative and who is not. I use innovative industries as a proxy for innovative establishments, but I cannot say whether individual establishments are innovative. Finally, some innovative industries may not appear innovative to all readers, for example, animal slaughtering facilities, however, many slaughtering establishments are innovative in order to improve productivity and sanitation (CREC, 2009).

4.2.4 Creating Entrepreneurial Industries

Entrepreneurship varies across space, making counties a suitably small unit of observation (Klein and Cook, 2006; Shrestha et al., 2007). Researchers can easily aggregate counties into labor market areas or metropolitan statistical areas. Most practitioners conduct economic development at the local level (Bartik, 1991; Wasylenko, 1997), so, when possible, studies of entrepreneurship should also be conducted at the local level.

Many argue that when examining change in entrepreneurship, the beginning and end points should coincide with business cycles (Chandra, 2002; Spelman, 2006). Examining entrepreneurship over a period of macroeconomic growth will lead to different results than examining entrepreneurship across a complete business cycle. Thus, I use the 2001-2006 period as the closest I can obtain to the 2001-2007 (peak-to-peak) business cycle. Regional business cycles do not necessarily coincide with national business cycles, but national business cycles are a reasonable proxy to use when conducting analysis for all counties in the county.[5] I cannot examine Entrepreneurial Industries over the 1991-2001 business cycle because NAICS was not established until 1997.[6]

To create the annual, county-level indicator of innovative entrepreneurship, *Entrepreneurial Industries*, I count the number of innovative industry establishments in both self employment data and single-unit employer establishment data. Self employment

[5] Atypical business cycles, generally caused by specific events, can influence entrepreneurship measures (Parker, 1996; Gartner and Shane, 2005). Hurricane Katrina and events on September 11th caused regional changes in employment and business activity.

[6] The years for which Entrepreneurial Industries are available is limited by both data availability and the implementation of NAICS, which replaced the SIC in 1997. Establishment birth data did not become NAICS-based until 1989-1999, and is only available through 2002-2003.

data are from the Census Bureau's Nonemployer Statistics and contain the number of establishments with no paid employees, e.g., proprietors, partnerships, in each county in each industry. Because these data are publically available, industries with less than three establishments are suppressed (see Appendix A). Single-unit employer establishment birth data are not publically available and were obtained through a special agreement with the Census Bureau; these data contain births in each industry in each county with no suppression. Counting innovative industry establishments in these data give me the number of innovative industry establishment births and self employed in each county, for each year. For instance, Champaign County, Illinois, had 19 innovative industry establishment births in 2003 and 334 employer establishment births in all sectors—thus the innovative industries establishment births represent 5.7 percent of establishment births in 2003. Similarly, Champaign County had 146 self employed in innovative industries in 2006, 1.30 percent of all self employed. Again, this is not a count of entrepreneurs, innovation, or innovative establishments; it is simply the method used to create Entrepreneurial Industries, an indicator of regional entrepreneurship.

Since counties are not homogenous, I must control turn the count into a rate (Audretsch and Fritsch, 1994; Gartner and Shane, 1995). For example, 19 births in Champaign County are meaningless without knowing the relative size of Champaign County. The choice of denominator can be a source of confusion and ambiguity because different methods of standardization lead to different results and conflicting policy signals (Audretsch and Fritsch, 1994). In this section, I discuss denominators for both self employment and establishment births and their appropriateness for Entrepreneurial Industries.

The theory of entrepreneurial choice explains individuals' entry into entrepreneurship and posits that *someone* starts each new business. Therefore, the denominator for self employment should represent everyone who could enter self employment (Evans and Jovanic, 1989). Total employment is a theoretically suitable denominator for self employment if we assume nascent entrepreneurs have some work experience. Many studies have adopted the same reasoning and use employment as the denominator for self employment because workers more closely represent the pool of nascent entrepreneurs than population or establishments (Iversen et al., 2008; Goetz and Rupasingha, 2008).

Despite theory suggesting employment as the denominator for self employment, I compare two denominators that could represent the pool of potential entrepreneurs, population and total employment. I measure population with BEA-REIS population estimates, because these data are available annually. I measure total employment with BEA-REIS total nonfarm employment. I exclude production agriculture employment because the Census data used to create Entrepreneurial Industries excludes it; the effect of this will be highest in rural areas and high-intensity agricultural areas such as California's Central Valley, where the Entrepreneurial Industries rate might be inflated because the denominator is smaller. Total jobs are a more accurate count of the pool of potential proprietors because it counts each job as equal; this is important because many proprietors are multiple-job holders. Total employment excludes the unemployed, but the advantages of including multiple job holding proprietors makes this tradeoff worthwhile.

The Entrepreneurial Industries self employment rate using population or total nonfarm employment are very similar (Table 4.4) and have a 0.928 correlation. Spatial analysis also points towards their similarity (Figure 4.3). Using population as a denominator results in increased heterogeneity, however (Figure 4.3). Heterogeneity makes some rates appear extreme, and increased heterogeneity makes statistical results less efficient. Thus, employment has both conceptual and empirical advantages over population as a denominator for Entrepreneurial Industries self employment.

Table 4.4 Entrepreneurial Industries Self Employment Variables

Variable	Mean	StDev	Definition
EI_se/emp	0.00155	0.00187	EI count for se divided by nonfarm total employment (BEA-REIS), 2000
EI_se/pop	0.00077	0.00093	EI count for se divided by population (BEA-REIS), 2000

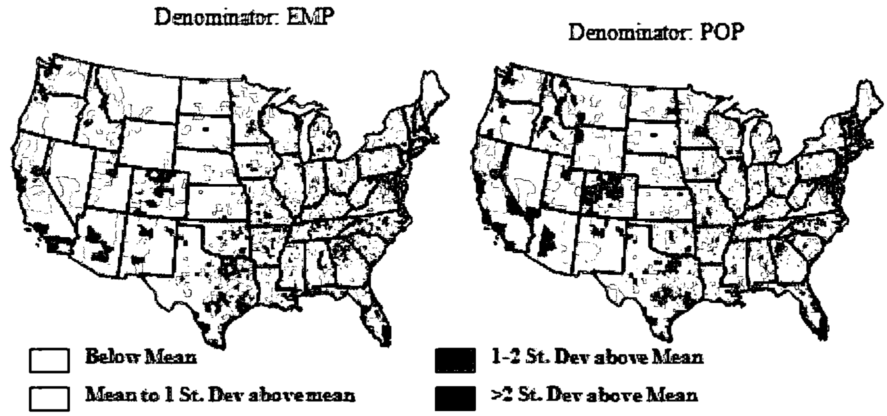

Figure 4.3 Entrepreneurial Industries Self Employment Using Nonfarm Employment and Population As Denominator

No established theory guides the choice of denominator for establishment births. An individual decides to enter self employment, but five individuals may partner to start an employer establishment. Thus, the theory of entrepreneurial choice cannot guide the choice of denominator for establishment births because the unit is not the individual.

Researchers have used establishments, population, labor force, and employment as the denominator for establishment births with no discussion of their appropriateness (Audretsch et al., 2002; Lee et al., 2004; Mueller, 2007). Others have found rates are affected by standardization approaches, implying that the selection of denominators affects results (Audretsch and Fritsch, 1994; Reynolds et al., 1994; Iversen et al., 2008). Given the implications of denominator choice, Love (1995) and Audretsch and Fritsch (1995), find the labor market, or people, approach is superior to establishment-based denominators for establishment births.

I normalize the count of Entrepreneurial Industries births using four variables—population, labor force, employment, and establishments—and compare the resulting rates.[7] Spatial analysis of the four rates shows that the population and labor force rates behave similarly (Figure 4.4). Indeed, the mean and standard deviation for the population rate is approximately half of that for the labor force rate, so we would expect similar maps based on their very similar distributions (Table 4.5). Both population and labor force control for the heterogeneity of counties, but labor force includes farm employment,

[7] Data are available annually at the county-level from BEA-REIS, BLS LAUS, and the Census Bureau's County Business Patterns, respectively.

so using population is more consistent with the establishment birth data, which exclude production agriculture establishments.

Although population is the preferred denominator for Entrepreneurial Industries, it is notable that establishments are too sensitive to the amount of entry that has already occurred to be a denominator for births (Audretsch and Fritsch, 1995; Love, 1995). Love (1995) found that using establishments as a denominator in an entrepreneurship model can produce the "wrong" signs in the model using an employment rate. Audretsch and Fritsch (1994) tested two classes of denominator, establishments and labor market. They found fault with the establishment rate because in areas with lots of small establishments, (potentially entrepreneurial areas), one additional birth makes little difference in the birth rate. In areas with a few relatively large establishments, one birth will dramatically increase the birth rate. As a result, two regions with the same population and same number of births can have vastly different birth/establishment rates if one region is dominated by small firms and the other dependent on a few large firms.

Very few counties had no innovative industries self employment in a given year, but many counties have no innovative industries births in a given year. Establishment births at the five-digit NAICS are relatively rare in all but the largest counties. As a result, one birth can make the rate appear unusually high. To overcome this problem, I use a 3-year moving average of birth counts in innovative industries.[8]

[8] To calculate the 3-year moving average for 2000, I sum innovative industry births for 99-00, 00-01, and 01-02 and divide by three to obtain an average, which I normalize by the base year population. The establishment birth data are available from 1998-1999 to 2002-2003.

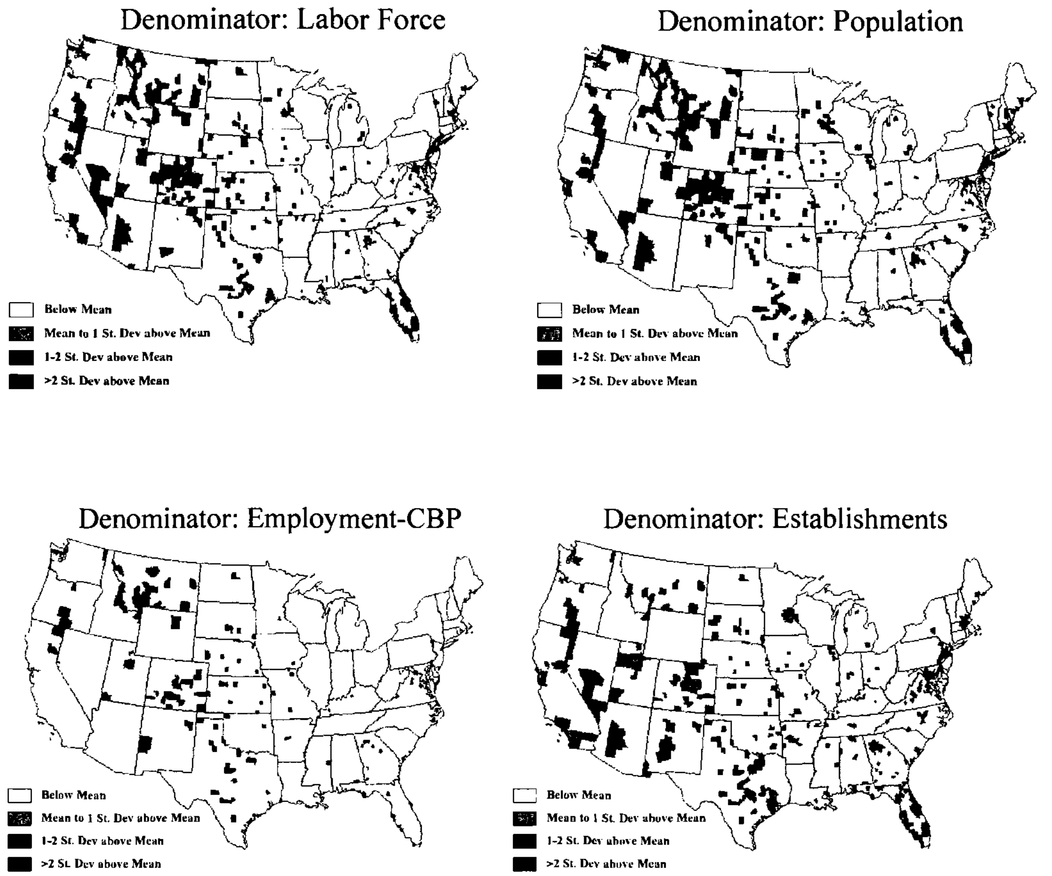

Figure 4.4 Entrepreneurial Industries Births (2000-2003) Using 2000 Denominator

Table 4.5 Summary Statistics Comparing Denominators

Variable	Mean	Std. Dev.	Definition
EI_birth/pop	0.00008	0.00007	EI count for births (Ave. of 1999-2001) divided by population, 2000 (BEA-REIS)
EI_birth/LF	0.000167	0.000139	EI count for births (Ave. of 1999-2001) divided by labor force, 2000 (BLS LAUS)
EI_birth/emp	0.000312	0.000528	EI count for births (Ave. of 1999-2001) divided by employment, 2000 (CBP)
EI_birth/estab	0.00336	0.00251	EI count for births (Ave. of 1999-2001) divided by establishments, 2000 (CBP)

4.3 ENTREPRENEURIAL INDUSTRIES RESULTS AND SENSITIVITY ANALYSIS

In this section, I establish that Entrepreneurial Industries is a conceptually and empirically valid indicator of entrepreneurship and is an improvement over other metrics because it considers innovation. Exploratory spatial data analysis and correlations suggest Entrepreneurial Industries has a positive relationship with growth and prosperity and differs from its parent measures. I also demonstrate the empirical validity of Entrepreneurial Industries by comparing it to widely publicized entrepreneurship indices available for some cities. Finally, I examine the sensitivity of Entrepreneurial Industries to the choice of innovative industries and find that Entrepreneurial Industries results are not sensitive to the choice of individual industries, suggesting Entrepreneurial Industries is robust to variation in innovative industries.

4.3.1 Construct Validity of Entrepreneurial Industries

4.3.1.1 Descriptive statistics

My two indicators of Entrepreneurial Industries, *EI_birth/pop* and *EI_se/emp* behave similarly spatially (Figure 4.5, 4.6). Both are highest in metro areas, notably Atlanta, Miami, Denver, Las Vegas, Dallas, Houston, and the San Jose/Silicon Valley/San Francisco metropolitan area, suggesting that the thick markets, access to labor, transportation, and intermediate goods in metro areas is associated with a higher level of Entrepreneurial Industries. Entrepreneurial Industries also appears high in regions with landscape and lifestyle amenities, such as Florida and the Rocky Mountains, suggesting high-amenity areas are associated with the level of Entrepreneurial Industries, e.g., footloose or lone-eagle entrepreneurs.

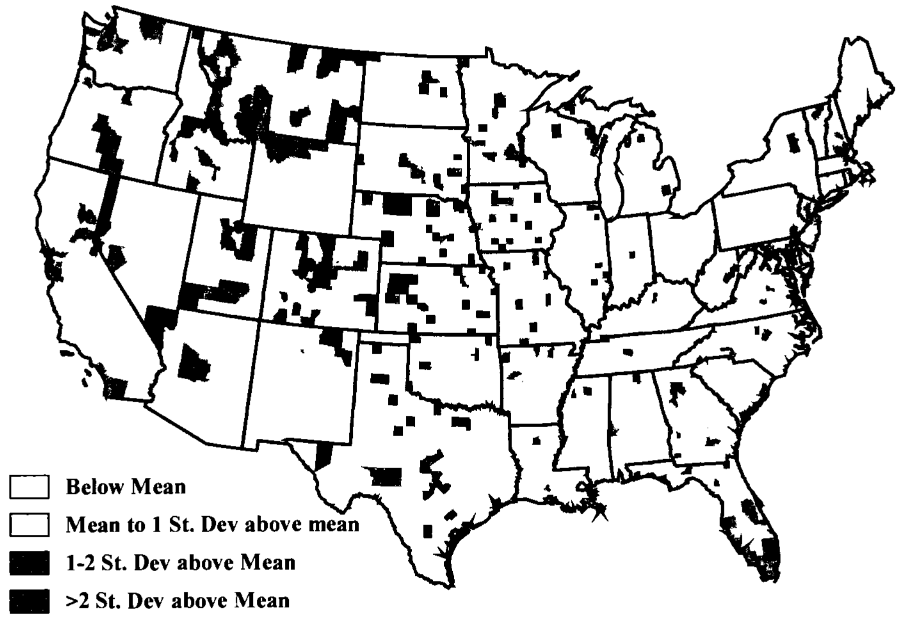

Figure 4.5 Entrepreneurial Industries Births, Three-Year Moving Average, 2000

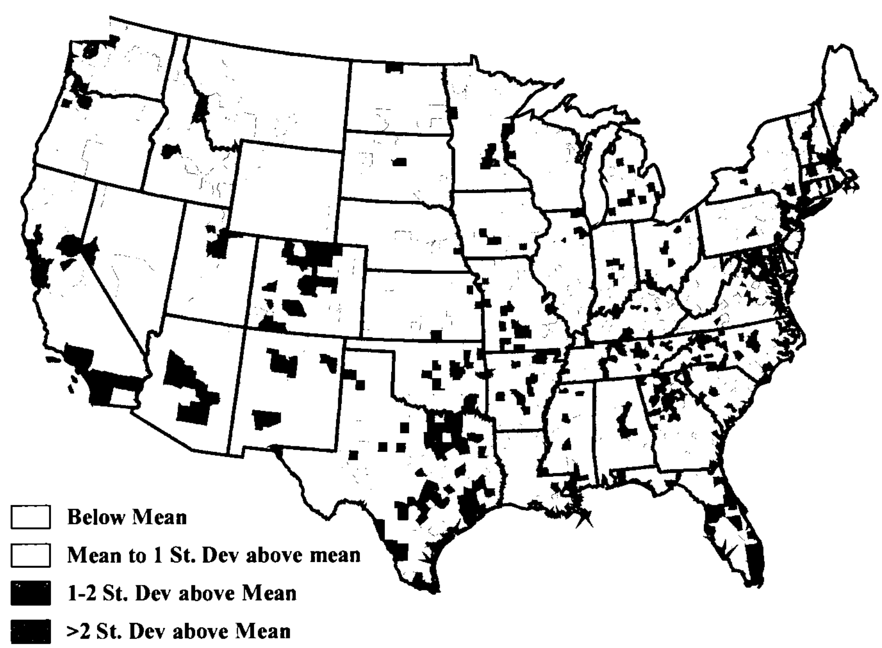

Figure 4.6 Entrepreneurial Industries Self Employment, 2000

Entrepreneurial Industries is also high in some non-metropolitan areas, for example, Cherry County, Nebraska, a large county in north central Nebraska. Cherry County had only two innovative industry establishment births in the three-year period but the low population makes the birth rate relatively high. Cherry County is surrounded by areas with no births, evidence that Entrepreneurial Industries births are sparse, even using a three-year moving average. The Cherry County case illustrates why I take the three-year moving average of Entrepreneurial Industries births. Using only one year would result in more unusually "high" rates in sparsely populated counties. Entrepreneurial Industries must be interpreted as a whole and over a suitable period of time, particularly in rural areas.

Entrepreneurial Industries may be low in sparsely populated areas because lower skill and lower technology occupations are more concentrated in rural areas (Massey 1984; Wojan 2000). Although rural areas have a higher proportion of high school graduates, they have a lower proportion of scientists, engineers, technicians, and other highly educated people than metro areas. Thus, they have fewer high skill industries.

EI_se/emp behaves differently from its parent measure, self employment (Figure 4.7). Spatial evidence suggests it is fundamentally different and includes the most innovative entrepreneurs and excludes most necessity-based self employment, implying the innovative industries method works to differentiate Entrepreneurial Industries from its parent measure (Figure 4.7). Figure 4.7 also shows Entrepreneurial Industries are highest in metro areas. *EI_se/emp* is highest in metro and amenity-driven areas, but self employment is highest in sparsely populated counties and most are necessity-based or lifestyle entrepreneurs who are self employed due to a lack of wage and salary job opportunities, rather than because they are innovating or creating something new (Henderson et al., 2006). I argue that the difference between *EI_se/emp* and its parent measure is that Entrepreneurial Industries identifies the most innovative entrepreneurs and excludes most necessity-based self employed.

EI Self Employment 2006

Self Employment 2006

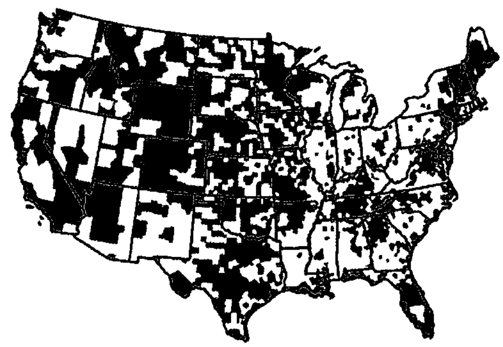

Figure 4.7 Entrepreneurial Industries Self Employment and Self Employment, Quartiles, Darker is Higher

Entrepreneurial Industries has a positive correlation with parent measures, *SelfEmp/Emp* and *Births/Pop*, but this correlation is not strong, further suggesting that notable differences between Entrepreneurial Industries and parent measures (Table 4.6). The correlation between *EI_se/emp* and its parent measure is 0.13 and 0.30 for *EI_birth/pop,* and suggests creating Entrepreneurial Industries was worthwhile empirically.

Positive correlation coefficients with employment growth, income growth, higher education, and Isserman's (2006) prosperity measure suggest Entrepreneurial Industries has construct validity (Table 4.6). Additionally, both Entrepreneurial Industries indicators exhibit a negative correlation with indicators of distress—the unemployment rate, poverty rate, and high school dropout rate. The Entrepreneurial Industries indicators have a positive correlation with McGranahan and Wojan's (2007) Recast Creative Class but the self employment and establishment birth rates have a negative correlation with Recast Creative Class (-0.17 and -0.05), illustrating how different Entrepreneurial Industries are from parent measures.

Table 4.6 Pearson Correlation Coefficients

Variable	EI_birth/pop	EI_se/emp	Description
EmpG01_06	0.201	0.360	Employment Growth 2001-2006, BEA-REIS
IncG01_06	0.135	0.187	Income Growth, 2001-2006, BEA-REIS
Unemp01	-0.158	-0.154	Unemployment rate, 2001, BLS LAUS
Poverty	-0.129	-0.249	Poverty rate, 2000, Decennial Census
Prosperity	0.078	0.114	Isserman's Prosperity (2005), 2000
RC Creative Class	0.244	0.453	McGranahan & Wojan (2007), 2000
Amenity Scale	0.163	0.212	McGranahan (1999) measure
Gartner: Estab/Pop	0.194	-0.096	Estabs (CBP, 2000) over Pop (BEA-REIS, 2000)
Births/Emp1000	0.300	0.128	Births, 2000 over 1000 employees (CBP, 2000)
SelfEmp/Emp	0.164	0.129	Nonemployers, 2000, over employment (CBP, 2000)
Patent/Pop	0.127	0.257	see text
Population	0.086	0.191	BEA-REIS, 2000
%CollegeEd	0.239	0.303	Percent >25 years with 4-year degree, Census, 2000
%HSdropout	-0.190	-0.152	Percent >25 years without HS or GED, Census, 2000

Using Isserman's (2006) Rural/Mixed Rural/Mixed Urban/Urban classification scheme, I find that the correlation signs discussed previously are consistent in both the most rural (*Rural*) and the most urban (*Urban*) counties (Table 4.7). The correlation coefficients for rural counties, however, are generally lower in magnitude. I do not include correlations for Mixed Rural and Mixed Urban for brevity.

In *Urban* counties, Entrepreneurial Industries is positively related to population. In *Rural* counties, the correlation between population and *EI_se/emp* is positive and higher than in *Urban* counties, but the correlation between population and *EI_birth/pop* is negative. This may be because self employment is more common in rural areas due to thin markets and a lack of wage and salary job opportunities (Low and Weiler, 2008). The negative correlation between Entrepreneurial Industries births and population in *Rural* counties is further evidence that employer establishment births are relatively uncommon in rural regions.

The correlation between Entrepreneurial Industries indicators and widely used measures of entrepreneurship (*Gartner*, *Births/Emp*, *SelfEmp/Emp*) suggests Entrepreneurial Industries is highest in urban areas (Table 4.7). This is likely because urban areas tend to have more patenting activity, more establishments, and more establishment births, even when normalizing for population (OhUallachain, 1999).

When interpreting differences in *EI_se/emp* across rural and urban areas, recall, *EI_se/emp* may undercount entrepreneurship in rural areas due to data suppression, but where innovative industries establishments are non-zero the rate may be inflated because the denominator excludes production agriculture employment. I need unsuppressed self employment data to explore the direction of the *EI_se/emp* bias in rural counties.

Table 4.7 Pearson Correlation Coefficients: Urban and Rural Counties

	Rural		Urban	
	EI_birth/pop	*EI_se/emp*	*EI_birth/pop*	*EI_se/emp*
EmpG01_06	0.10	0.24	0.31	0.33
IncG01_06	0.07	0.09	0.22	0.16
Unemp01	-0.15	-0.07	-0.25	-0.06
Poverty	-0.07	-0.15	-0.39	-0.25
Prosperity	0.06	0.05	0.21	0.12
RC Creative Class	0.16	0.34	0.49	0.27
Amenity Scale	0.11	0.17	0.34	0.29
Gartner: Estab/Pop	0.20	-0.11	0.21	-0.17
Births/Emp1000	0.28	0.13	0.60	0.67
SelfEmp/Emp	0.18	0.19	0.37	0.67
Patent/Pop	0.02	0.11	0.38	0.20
Population	-0.13	0.41	0.10	0.18
%CollegeEd	0.21	0.13	0.39	0.18
%HSdropout	-0.17	-0.03	-0.31	-0.07

4.3.1.2 Empirical comparison of entrepreneurial industries and other indicators

County-level entrepreneurship indices are rare. Most assess entrepreneurship in metropolitan areas only. Seven metro areas had at least one county in the top 50 *EI_se/emp* and top 50 of *EI_birth/pop*; indeed, all seven metros had multiple counties within the top 50. These metro areas are Atlanta, Dallas, Denver, Miami, New York City, San Francisco, and Washington, DC (Table 4.8). Although a crude substitute for relevant metro rates of Entrepreneurial Industries, my list is comparable to others' indexes that examine metro areas.

Entrepreneurial Industries is consistent with the most recent and best-known metro entrepreneurship index, the Kauffman Index of Entrepreneurial Activity (Fairlie, 2009). The 2008 index is complex and computed using Current Population Survey data on self employment and employer establishment births. All top Entrepreneurial Industries metro areas are in the Kauffman top ten, except Denver (Table 4.8). Some large cities do not appear in either the Kauffman top ten or the Entrepreneurial Industries top seven

(Chicago, Philadelphia, Boston, Detroit, and Seattle), suggesting that city size does not predict current entrepreneurial activity.

Table 4.8 Kauffman Top Ten and Entrepreneurial Industries Cities

Kauffman Top 10	EI Top 7
Atlanta	Atlanta
Phoenix	
Riverside, CA	
Los Angeles	
Miami	Miami
New York City	New York City
San Francisco	San Francisco
Dallas	Dallas
Houston	
Washington, DC	Washington, DC
	Denver

Further empirical support can be gleaned from older entrepreneurship indexes. *Inc.* Magazine's Top Entrepreneurial Cities (1990) listed Las Vegas as the most entrepreneurial city, with the top ten cities including Washington, Orlando, Tallahassee, San Jose, Atlanta, Charleston, SC, Lincoln, NE, Raleigh-Durham, NC, and Anaheim, CA (Case, 1990). Although only Washington and Atlanta are top Entrepreneurial Industries metros, Figures 4.5 and 4.6 show Entrepreneurial Industries is high in Las Vegas (southern Nevada), Florida, and parts of the south, including Atlanta, Charleston, Charlotte and the Research Triangle. Thus, *Inc.*'s top ten differ from the top seven Entrepreneurial Industries metros, but their top metros are all in Entrepreneurial Industries' top quartile.

Finally, *Entrepreneur* Magazine's Best Large Cities for Entrepreneurship, 2006, are, in-order, Phoenix, Charlotte, Research Triangle, NC, Las Vegas, Austin, Washington, DC, Memphis, Nashville, Norfolk/Virginia Beach, and San Antonio. Only one of the seven Entrepreneurial Industries cities is on this list but all ten are high for either Entrepreneurial Industries self employment or Entrepreneurial Industries births, suggesting the top Entrepreneurial Industries metros are similar to other indices' top entrepreneurial metropolitan areas. For example, Phoenix has shown up on several of the entrepreneurial indexes and is not among the high Entrepreneurial Industries cities; it has a high level of Entrepreneurial Industries births but not Entrepreneurial Industries self employment.

4.3.2 Variations for Sensitivity Analysis

To differentiate Entrepreneurial Industries from widely used entrepreneurship measures, I developed a method for identifying innovative industries, since data on innovative establishments are not available. The cost of identifying innovative entrepreneurs via innovative industries is that I must *estimate* the number of innovative establishments by defining some industries as innovative and others, not. In this section, using two alternative methods, I assess the sensitivity of Entrepreneurial Industries to the selection of innovative industries. The inclusion or exclusion of individual industries does not affect Entrepreneurial Industries results. Focusing on the inclusion/exclusion of specific industries, thus, is futile.

The Entrepreneurial Industries method, ST3, requires innovative industries have three times the mean level of high skill or high tech employment, a demanding criterion that less than 40 industries met. Two less demanding alternatives, ST2 and STP, test how sensitive Entrepreneurial Industries is to the inclusion of specific industries.

The ST2 method lowers the cutoff required to meet the high skill and high tech primary criteria to two times the mean, hence ST2, and keeps the secondary criterion requirement. Two times the mean is the lowest threshold for high tech industries, as defined by the BLS (its "low content" high tech industries).

Using the ST2 method, the number of innovative industries rose from 39 to 61 and included more service industries (NAICS 51-81). Some newly included industries are Casino Hotels (NAICS 72112), Independent Artists, Writers, and Performers (NAICS 71151), Services for the Elderly and Persons with Disabilities (62412), and Monetary Authorities-Central Banks (NAICS 52111). The owner/operator and risk/uncertainty bearing attributes of entrepreneurship will screen out some of these innovative industries from the Entrepreneurial Industries indicators, e.g., the Federal Reserve bank system (NAICS 52111) has no self employed, and has had no births in the last 15 years. As a result, Fed establishments would not appear in either the self employment or birth data.

Recall, with ST3 no industries met both high skill and high tech; there was considerable overlap between high skill and high tech using ST2, mainly in the service sectors.

The STP method differs from ST3 and ST2 by forgoing the requirement for a secondary criterion but allowing high patent to stand alone as a primary criterion. Although the patent data are not ideal due to data problems, discussed in Section 4.1, industries with high levels of patenting are at the extreme of innovation—so much so that firms are willing to spend the time and money necessary to patent the new technology, and presumably reap sizeable economic rents from these patents.

Using the STP method, the number of innovative industries rose from 39 to 70—more industries than even the ST2 method includes (Table 4.9). Adding patents as a primary criterion, rather than eliminating the secondary criterion, accounted for most of this increase and the majority of industries new in STP are manufacturing industries (Table 4.9). Indeed, most of the innovative industries meet more than one secondary criterion, suggesting the principal hurdle is the high standard necessitated by meeting three times the mean for high skill or high tech.

Table 4.9 Count of Industries Meeting ST3, ST2 or STP at 3-Digit NAICS

Naics	Description	ST3	ST2	STP
221	Utilities	1	2	2
237	Heavy and Civil Engineering Construction		1	
311	Food Manufacturing	1	1	2
325	Chemical Manufacturing	2	5	7
326	Plastic & Rubber Product Mfg.			1
331	Primary Metal Manufacturing			2
332	Fabricated Metal Product Manufacturing		1	1
333	Machinery Manufacturing		1	2
334	Computer and Electronic Manufacturing	6	6	6
335	Electrical Equipment Manufacturing		1	1
336	Transportation Equipment Manufacturing	2	2	4
337	Furniture and Related Product Manufacturing	1	2	3
423	Merchant Wholesalers, Durable Goods	1	2	1
441	Motor Vehicle and Parts Dealers	1	1	1
486	Pipeline Transportation		1	
511	Publishing Industries (except internet)	1	1	1
516	Internet Publishing and Broadcasting	1	1	1
517	Telecommunications	2	3	2
518	Internet Service Providers	2	2	2
521	Monetary Authorities-Central Bank		1	
541	Miscellaneous Professional, Scientific, and Technical Services	6	8	6
551	Management of Companies and Enterprises	1	1	1
561	Administrative and Support Services	5	5	8
562	Waste Management and Remediation Services		1	
621	Ambulatory Health Care Services	1	2	1
622	Hospitals	1	2	1
623	Other Residential Care Facilities			1
624	Nursing and Residential Care Facilities	1	3	1
711	Performing Arts, Spectator Sports, and Related Industries		1	
713	Amusement, Gambling, and Recreation Industries		1	
721	Accomodation		1	
811	Repair and Maintenance	1	1	1

Summary statistics show the alternative methods have higher means than Entrepreneurial Industries (ST3), because they include more industries, which makes the count higher (Table 4.10). The mean for births is an order of magnitude lower than the mean for self employment. The mean for *STP* applied to self employment is much higher than the mean for *STP* applied to births, and its standard deviation is much higher too, suggesting there is much more variation in STP than ST2 or ST3, which could be problematic for use in entrepreneurship analysis.

Table 4.10 Summary Statistics for Alternate Methodologies

	Obs.	Mean	St. Dev.	Min.	Max.
ST3_birth	2635	0.00019	0.00020	0	0.0026
ST2_birth	2777	0.00030	0.00026	0	0.0039
STP_birth	2762	0.00024	0.00022	0	0.0028
ST3_SE	3068	0.0034	0.0022	0	0.0190
ST2_SE	3077	0.0083	0.0050	0	0.0396
STP_SE	3076	0.01164	0.0067	0	0.0745

I map the three indicators, based on their distribution to examine their spatial distribution (Figure 4.8). STP is visually different from ST2 and ST3. STP is high in southern Appalachia from the Atlanta metro area to the Ozarks. This is likely due to a concentration of independent manufacturing in this region. Compared to ST3, the ST2 method results in more counties at the high end of the range (greater than two standard deviation above average), noticeably so in the front range of the Rocky Mountains and in the Northeast. Otherwise, ST2 and ST3 are spatially similar. I prefer ST3 to ST2 though, because it has a higher innovation criterion.

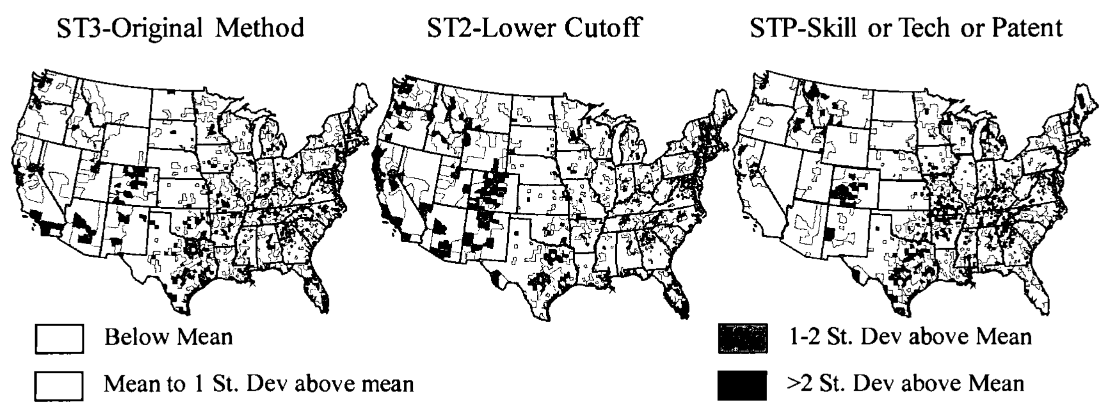

Figure 4.8 Comparison Of Indicators Using Three Methodologies

4.4 CONCLUSION

Entrepreneurial Industries is conceptually valid, empirically valid, and robust to the selection of innovative industries. I demonstrated construct validity by showing Entrepreneurial Industries results are correlated with widely used entrepreneurship indices and that Entrepreneurial Industries is positively correlated to growth and prosperity and negatively with indicators of economic distress. Entrepreneurial Industries represents an improvement over other entrepreneurship measures and indicators because it captures multiple dimensions of entrepreneurship, including innovation. Entrepreneurial Industries is substantively different from its parent measures. Finally, Entrepreneurial Industries is available annually for U.S. counties, allowing it to serve as a useful building block for regional analysis across space and time.

Entrepreneurial Industries has the potential to improve regional research and economic development practice and policymaking, making policies and programs more effective and less costly. Additionally, using an entrepreneurship benchmark that excludes lifestyle or necessity-based entrepreneurs may alter perceptions of regional entrepreneurship and highlight programmatic needs and successes.

I cannot identify innovative establishments, so I proxy with the most innovative industries and argue that their nexus with establishment births and the self employed is a useful *indicator* of entrepreneurship. It is not a count of entrepreneurs. Although such a count might be ideal, it is unobtainable on an annual basis for U.S. counties. To differentiate Entrepreneurial Industries from other entrepreneurship measures I employ high standards for defining innovative industries at the cost of not including some innovative establishments and including some non-innovative establishments. Sensitivity analysis, however, shows Entrepreneurial Industries results are robust to variation in the choice of innovative industries, thus, arguing about the inclusion of a specific industry will not change the results.

Entrepreneurial Industries in rural counties require careful interpretation and additional data could improve Entrepreneurial Industries in rural areas. Unsuppressed self employment data would improve Entrepreneurial Industries, particularly in sparsely populated counties where one or two establishments, now disclosed, could lead to a high Entrepreneurial Industries rate. Additionally, rural areas suffer from construct validity problems where a smaller stock of entrepreneurs exists and the denominator, employment, is shrinking, which can make Entrepreneurial Industries appear to increase over time (McGranahan and Wojan, 2007).

The list of innovative industries is only an indicator of innovativeness; it is not definitive, and should not be used for industry targeting. Additionally, Entrepreneurial Industries is static over both space and time due to its construction, preventing analysis of the spread of innovation across space or the change in high skill/high tech employment over time. Being able to capture change in innovation across space would enable me to define innovative industries better and change in innovation over time would also help, by identifying where innovation and/or automation is affecting the percent high skill and high tech employment. Finally, data on Entrepreneurial Industries employment, productivity, or value added could improve the indicator and the definition of innovative industries, but to conduct such analysis I need micro data from the Center for Economic Studies.

CHAPTER 5: ENTREPRENEURIAL INDUSTRIES: ENTREPRENEURSHIP MODEL

While much of the research on the determinants of entrepreneurship is at the national level (Acs and Armington, 2006), research on the regional determinants of entrepreneurship is growing (Gebremariam et al., 2006; Goetz and Rupasingha, 2008). This interest has emerged from research that suggests a strong connection between entrepreneurship and growth. Many studies use measures of entrepreneurship that ignore innovation, such as self employment, despite innovation being a long established defining attribute of entrepreneurship (Schumpeter and Opie, 1983). This chapter examines the determinants of Entrepreneurial Industries and contributes to the determinants of entrepreneurship literature.

Evans and Leighton (1989) conducted one of the first studies on the determinants of self employment, using longitudinal micro data on white males who identified themselves as self employed. Parker (1996) and Schiller and Crewson (1997) built upon this initial work by broadening the sample and providing a theoretical foundation for the model. This research has broader applicability as research that is more recent showed that determinants of firm births resemble the determinants of self employment (Lee et al., 2004; Acs et al., 2006). Common determinants include entrepreneurial attitudes (fear of failure, goal-setting, confidence in abilities), access to capital, firm characteristics, and geographical environment.

Recent research examining the determinants of *regional* self employment has found region-specific factors affect entrepreneurship (Georgellis and Wall, 2000). Most regional entrepreneurship research, however, focuses on urban regions, and often omits rural places. Glaeser (2007) found that, in metro areas, self employment rates were highest for men and rise with age, educational attainment, and population. Little is known about the determinants of self employment in nonmetro areas though. Goetz and Rupasingha (2008) use all U.S. counties in their entrepreneurship model and find a nonmetropolitan binary variable is significant, indicating the existence of a rural/urban difference in entrepreneurship levels.

In this chapter, I examine the drivers of Entrepreneurial Industries, which is unique because it incorporates innovation. I build on Goetz and Rupasingha's (2008) county-level entrepreneurship model and find that the drivers of innovative entrepreneurship differ from self employment. The entrepreneurship model explains more variation in Entrepreneurial Industries growth than self employment growth suggesting factors can explain Entrepreneurial Industries more than self employment. I find financial collateral and amenities positively influence Entrepreneurial Industries, while self employment is driven by a negative relationship with income and income growth. Results suggest entrepreneurship models are sensitive to the entrepreneurship measure and that Entrepreneurial Industries may be more useful to policymakers and economic development practitioners who would like to promote innovative entrepreneurship, rather than necessity-based entrepreneurship in their regions.

5.1 DATA

I use Goetz and Rupasingha's (2008) model as a starting-point for my model because it is the state-of-the-art model and incorporates the findings of other regional entrepreneurship models. Explanatory variables are drawn from literature on the determinants of entrepreneurship, and address individual and regional characteristics associated with entrepreneurial activity. These explanatory variables also reflect findings of prior work on modeling regional self employment and establishment birth rates (Evans and Leighton, 1989; Parker, 1996; Lee et al., 2004).

5.1.1 Independent Variables

Like other models (Lee et al., 2004; Acs et al., 2006), Goetz and Rupasingha's dependent variable is expressed as a function of demographic, Ω, regional, Ψ, and policy variables, Γ, Equation (5.1).

$$y_i = f(\Omega_i, \Psi_i, \Gamma_i) = f(\beta \mathbf{X}_i + \varepsilon_i) \tag{5.1}$$

Goetz and Rupasingha (2008) use lagged explanatory variables to reduce the endogeneity bias and show that at least some of the causality flows from explanatory variables to self employment. Goetz and Rupasingha examine growth in nonfarm proprietorships between 1990 and 2000, but I use 2000 as the base year and examine

Entrepreneurial Industries over the most recent business cycle, in part, because I found unusually robust growth during the 1990s affected test results.

5.1.1.1 Demographic variables Ω

In Equation (5.1), Ω represents collateral, human capital, and other demographic characteristics. Collateral facilitates borrowing capital and represents the ability of nascent entrepreneurs to obtain financing for entrepreneurial ventures; thus, I expect to find a positive coefficient (Goetz and Freshwater, 2001). A county's median home value, *HomeValue,* suggests the overall level of collateral available for a loan, while the rate of owner occupied homes, *HomeOwn,* gives the share of people who have the collateral available to them (Table 5.1).

Measures of high school and college educational attainment are included in the model to control for human capital's role in determining entrepreneurship rates, and I expect this relationship to be positive (Evans and Leighton, 1989; Audretsch and Fritsch, 2002). I control for human capital with the percent of adults, over age 25, who are college educated, *College,* and percent of adults who graduated from high school or receive a GED, but not college, *HS.* I use both because I am interested in the hypothesis that education has a U-shape relationship with entrepreneurship (Goetz and Rupasingha, 2008).

Goetz and Rupasingha include age, ethnicity, and gender because research suggests entrepreneurs are more likely to be male and older than the general population (Goetz and Freshwater, 2001). To control for these relationships, the model includes the percent of the population that is Caucasian, *White,* the county median age, *MedAge*, and the female percentage of the labor force, *Female*.

5.1.1.2 Regional variables, Ψ

Financial capital, labor market, economic structure, and other place-based characteristics are represented by Ψ in Equation (5.1). Local bank deposits per capita give insight on the region's availability of financial capital, particularly in rural areas where venture capital does not usually exist (Garofolli, 1994; Low et al., 2005). *DeposPop* measures how much money local banks have on hand for small business loans and, although loan decisions are increasingly made with score-carding or at distant headquarters, this variable is useful in assessing financial capital's availability on past entrepreneurship. The Federal Deposit Insurance Corporation make these data available annually for all counties, and this is one of the few county-level datasets on the availability of financial capital (Table 5.1).

The unemployment rate, *Unemp*, and its square, *UnempSq*, are widely used in growth and entrepreneurship models; I include both due to expected nonlinearities in the coefficient. Parker (1996) hypothesizes that low unemployment "pulls" people into entrepreneurship due to the buoyant regional economy and high probability of success. Conversely, high unemployment can also lead to entrepreneurship because people are "pushed" into entrepreneurship due to a lack of wage and salary job opportunities, e.g., the jobless recoveries following the 1991 and 2001 recessions.

I use average wage and salary income, *WSinc,* as a proxy for available income. I argue average wage and salary income represents the opportunity cost of leaving wage and salary employment to enter self employment or start a new business—the tradeoff many nascent entrepreneurs face (Low and Weiler, 2008). I also include growth in wage and salary income, *WSincGro,* because the level is affected by past economic conditions and the change reflects current economic conditions, which affect the decision to enter self employment.

I include a set of industry employment variables after ensuring there is no collinearity between industry employment and the Entrepreneurial Industries indicators, although both are based on industries. I include percent employment in *Ag*, agriculture and forestry, *Mining*, *NonDurManu*, non-durables manufacturing, *DurManu*, durables manufacturing, *Trade*, wholesale and retail trade, *Visitor*, recreation, arts, accommodation, and food services, and *Services*, information, finance, insurance, real estate, and professional, scientific, and technical services. Other industries are the omitted condition.

Entrepreneurial Industries is higher in metropolitan areas and amenity-driven places, so I control for both. I include a dummy variable, *Nonmetro,* to test whether the nonmetro status of a county affects entrepreneurship and control for this expected relationship. I control for the attractiveness of place using McGranahan's (1999) amenity index, *Amenity*. The index includes measures of topography, weather, and water, and I expect it to be positive because others have found footloose entrepreneurs are attracted to regions with natural and scenic amenities. Additionally, Entrepreneurial Industries is high in amenity-driven regions such as Florida and the Rocky Mountains.

Finally, I include the level of the dependent variable in the growth model, Equation 5.9, to control for the existing base of entrepreneurs. Goetz and Rupasingha include variables on the relative risk and return of their dependent variable, self employment. Data on the risk and return of Entrepreneurial Industries are not available, publically, for counties, and the relative risk of Entrepreneurial Industries is a part of the definition (churn), so I do not include variables on the relative risk and return of Entrepreneurial Industries in the model.

5.1.1.3 Policy variables, Γ

I use a state income tax index to represent the policy vector, Γ, in this model because the majority of self employed and establishment births revenue flow through individual income taxes and comparing state tax policies other than on income is difficult. The Individual Income Tax index, *Tax,* is from the Tax Foundation and is an index with a scale of zero to ten; zero being the worst and 10 being the best. The Tax Foundation's background paper on tax indices contains more, detailed information on this variable (Barro, 2008).

I think Goetz and Rupasingha's state-level policy vector is too aggregated to be meaningful and makes interpreting the coefficient difficult.[9] Goetz and Rupasingha use state-level indices of economic freedom, which measure the size of government, taxation, and labor market freedom. I think the Individual Income Tax index is more relevant for

---

[9] I do not use the Business Tax Climate Index, which is used in several other determinants studies, because all self employed businesses' revenue flows through individual income taxes and the majority of establishment births flow through individual income taxes (S-Corporation, Partnerships, and Sole Proprietorships are all taxed via individual income taxes, leaving only C-corporations). I do not know the legal form of establishments in my birth data, but the Brookings Institution reports that only 7.1 percent of small business returns are from C-corps. Thus, if 92.1 percent of small businesses have their revenue taxed via individual taxes, then the Individual Income Tax Index may be more appropriate than the Business Tax Index. Higher taxes increase entrepreneurship because the potential to evade taxes is higher (Parker, 1996); however, high taxes may reduce entrepreneurship because of higher personal and self employment taxes, therefore, the expectation of the coefficient sign is ambiguous.

the self employed and new/small firms because income tax affects all firms but government size, regulations, and labor policies, in aggregate, are marginally relevant.

Table 5.1 contains all the explanatory variables used in my model, their description, source, and summary statistics. I also include whether the same variable was used by Goetz and Rupasingha, or if not, what variable was used in Goetz and Rupasingha's model.

Table 5.1 Explanatory Variables for Entrepreneurship Model

Goetz & Rupasingha		Explanatory Variables	Source	Mean	StD	Min	Max
same Ω	*HomeValue*	% of residences which are owner occupied	#	82806.1	46280.5	0	1E+06
same	*HomeOwn*	Median home value ($)	#	0.742	0.073	0.196	0.899
same	*College*	% >25 years with a 4-year college degree	#	0.163	0.075	0.049	0.605
same	*HS*	% >25 years high school diploma, highest degree	#	0.349	0.065	0.109	0.532
same	*MedAge*	Median age	#	37.3	3.9	20.6	54.3
same	*Female*	% female	#	0.458	0.021	0.231	0.541
same	*White*	% white	#	0.851	0.159	0.050	1.000
PCI Ψ	*WSinc*	Average wage and salary income	*	20.6	5.6	0.0	68.6
Change in PCI	*WSincGro*	Average wage and salary income	*	0.493	0.382	-0.527	11.086
same	*DeposPop*	Bank deposits ($1000) per capita	FDIC	0.012	0.014	0.000	0.480
same	*Unemp*	% unemployed	#	0.048	0.026	0.000	0.277
same	*UnempSq*	Unemp squared	#	0.003	0.004	0.000	0.077
same	*Nonmetro*	non-metropolitan	OMB	0.662	0.473	0.000	1.000
same	*Amenity*	McGranahan's amenity scale	see text	0.056	2.316	-6.400	11.170
same	*Ag*	% of employed in agriculture & forestry	#	0.063	0.071	0.000	0.556
same	*Mining*	% in mining	#	0.012	0.027	0.000	0.456
same	*NonDurManu*	% in nondurable manufacturing	#	0.071	0.051	0.000	0.431
same	*DurManu*	% in durable manufacturing	#	0.089	0.064	0.000	0.420
same	*Trade*	% in retail and wholesale trade	#	0.145	0.025	0.017	0.299
same	*Visitor*	% in arts, recreation, food, and accommodation	#	0.106	0.044	0.000	0.411
same	*Services*	% in information, FIRE, Prof & Sci Services	#	0.071	0.033	0.000	0.364
Economic Γ freedom index	*Tax*	Individual Income Tax index, FY July 2005	see text	5.8	2.0	2.0	10.0

All data are 2000, unless otherwise noted

* Bureau of Economic Analysis, Regional Economic Information System (REIS)

\# Decennial Census of Population, 2000

5.1.2 Dependent Variables

I examine both the level of entrepreneurship and entrepreneurship growth over the most recent business cycle using Entrepreneurial Industries. I examine growth between 2001 and 2006 because data for Entrepreneurial Industries are not available prior to 1997, precluding analysis of the 1991-2001 business cycle. Descriptive statistics for all dependent variables are in Table 5.2.

Data to calculate change in births are not available, so I evaluate the determinants of the level of Entrepreneurial Industries using the Entrepreneurial Industries indicators discussed in Chapter 4, *EI_birth/pop* and *EI_se/emp,* as well as *Prop*, the nonfarm proprietorship, or self employment, rate for discussion purposes.

To examine the determinants of growth in entrepreneurship, my dependent variables are *EI_se_chg* and *Prop_chg,* which are the change, or growth, in *EI_se/emp* and *Prop* (Table 5.2). Goetz and Rupasingha calculate their dependent variable, the proprietor growth rate, as the proprietorship rate at time $t+1$ minus the proprietorship rate at time t (Equation 5.2), and I calculate my dependent variables the same way (Equation 5.3).

Table 5.2 Dependent Variables

		Dependent Variables	*Mean*	*StDev*	*Min*	*Max*
Y	*EI_se/emp*	EI applied to self employment data" /nonfarm employment*	0.0034	0.0022	0.0000	0.0190
Y	*EI_birth/pop*	EI applied to births% /population*, 3-year MA 1999-2001	0.00008	0.00007	0	0.00071
Y	*Prop*	nonfarm proprietor employment*/total nonfarm employment*	0.247	0.093	0.030	0.710
$Y_{t-(t-1)}$	*EI_se_chg*	EI_se/emp_{2006}-EI_se/emp_{2001}	0.0019	0.0016	-0.0043	0.0148
$Y_{t-(t-1)}$	*Prop_chg*	$Prop_{2006}$-$Prop_{2001}$	0.0316	0.0394	-0.2093	0.3976
Y	*ST2*	ST2 industries applied to self employment data" /nonfarm emp*	0.008	0.005	0	0.040
Y	*STP*	STP industries applied to self employment data" /nonfarm emp*	0.012	0.007	0	0.075

* Bureau of Economic Analysis, Regional Economic Information System (REIS)
" Bureau of Census, Nonemployer Statistics, 2006, unless otherwise noted

$$(5.2)\quad \Delta proprietor = prop_{00} / totemp_{00} - prop_{90} / totemp_{90}$$

$$(5.3)\quad \Delta y = EI\_se_{2006} - EI\_se_{2001} = EI\_se / totemp_{2006} - EI\_se / totemp_{2001}$$

5.2 MODEL

In initial OLS estimations, I find no evidence of multicollinearity but do find evidence of heteroskedasticity. The Breusch-Pagan (BP) test for heteroskedasticity rejects the null hypothesis of no heteroskedasticity in the error terms, BP=575.06 and $p < 0.0001$. Heteroskedasticity in the OLS model is one of the first indicators that the errors contain a spatial process. Based upon visual heteroskedasticity in the map of the dependent variables (Figure 4.5, 4.6) and the map of the OLS residuals (Figure 5.1), it appears that spatial processes may be driving the heteroskedasticity. The Moran's I, a test statistic for spatial autocorrelation, is positive and significant (p=0.019), indicating spatial processes in OLS residuals.

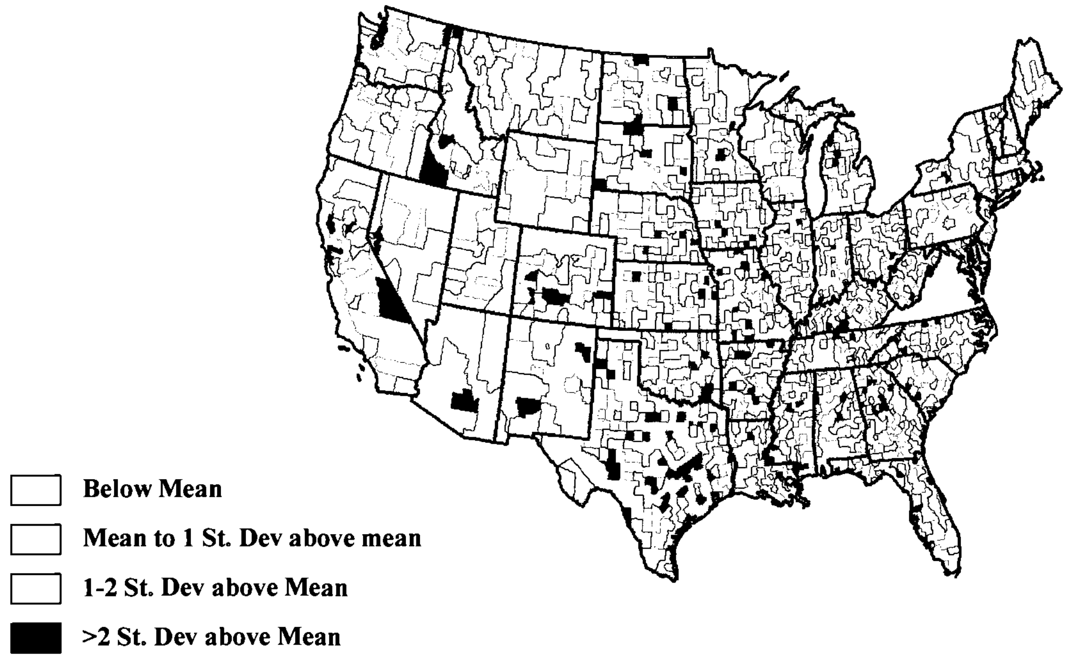

Figure 5.1 OLS Residuals*

*Virginia excluded due to missing data

5.2.1 Spatial Econometric Model Specification

Spatial processes are common in U.S. county-level models and more resent research usually attempts to control for it using spatial econometrics, including Goetz and Rupasingha (2008). They, however, incorrectly specify their spatial econometric model and interpret non-identified coefficients. I correct this problem and estimate the correct model.

Goetz and Rupasingha use the General Spatial Model (SAC) spatial econometric model that incorporates both spatial error and spatial lag terms (Equation 5.4). Employing such a model, however, often leads to identification problems and should be avoided unless strong theoretical reasons exist (Florax and Rey, 1995). [10] Both nuisance (error)

[10] Detecting the presence of both spatial error and spatial lag processes is difficult because the LM test tends to be significant when either the error or the lag alternative hypothesis is proper, but not necessarily both, due to the specified null hypothesis (Anselin, 2008b). The LM test with alternative hypothesis of a higher order alternative model, with both a spatial error and a spatial lag term is possible; however, rejection of the null of this test does not necessarily imply that the higher order model is the proper alternative. In many cases, re-specification of the spatial W matrix can change the LM test results.

errors and substantive (lag) errors exist in most U.S. county-level models, but one dominates the other and only the dominant type of spatial dependence should be controlled for with the appropriate model, e.g., Spatial Error Model (SEM) (see Appendix B, Equation B.1) or the Spatial Autoregressive Model SAR (Appendix B, Equation B.2). Higher-order models, like SAC, attempt to control for both the nuisance (error) and substantive (lag) dependence. Higher-order spatial models, however, can lead to identification problems that can be controlled for by using either the lag or error model, but not both.

I illustrate the problem with Goetz and Rupasingha's model. In time-series analysis, the SAC model, Equation 5.4, is similar to a first-order autoregressive model with serially correlated errors. The SAC spatial model is much more complex, however, and requires great care to ensure proper identification (Anselin, 2008a). If $W_1 = W_2$ or the spatial weights are not correctly specified the weights matrix is in both the error term and an explanatory variable—creating a substantial identification problem (Anselin, 2008b).

(5.4a) $Y = \rho W_1 Y + \beta X + \varepsilon$, where

(5.4b) $\varepsilon = \lambda W_2 \varepsilon + u$, with $\mu \sim N(0, \sigma^2 I_n)$.

Following Anselin (2008a), I rewrite the SAC model to illustrate the identification problem:

(5.5) $y = \rho W_1 y + \lambda W_2 y - \rho\lambda W_2 W_1 y + X\beta - \lambda W_2 X\beta + \mu$.

If $W_1 W_2$ are non-overlapping ($W_1 W_2$=0), we have:

(5.6) $y = \rho W_1 y + \lambda W_2 y + X\beta - \lambda W_2 X\beta + \mu$.

In practice, however, the same W is often used. Goetz and Rupasingha use the same W matrix for both, the k=3 nearest neighbors matrix. Thus, $W_1 = W_2$. Rearranging:

(5.7) $y = (\rho + \lambda)Wy - \rho\lambda W^2 y + X\beta - \lambda WX\beta + \mu$

When β =0 this model, Equation 5.7, is not identified (Kelejian and Prucha, 1998). Goetz and Rupasingha have some zero coefficients, which results in the entanglement of rho and lambda (Anselin, 2008a).

To address the identification problem, the weights matrices could be re-specified (Wojan et al., 2007), or a procedure for interpreting LM tests should be followed (Florax and Rey, 1995; Appendix B.4). If Florax and Rey's (1995) method is adopted, the dominant type of spatial dependence is controlled for and the identification problems and W specification problems are avoided. Wojan et al. (2007) do not follow Florax and Rey's procedure, but address the problem by using social weights and geographic weights to specify $W_1 \neq W_2$ in a higher-order model, circumventing the non-overlapping weights problem, but requiring novel solutions to define the same neighbors in different ways. This approach is rarely used because the parameters in social/spatial interaction models are identified only under strict conditions (Manski, 1993; Anselin 2008a) and a mis-specified W matrix could change the alternative hypothesis of the LM test (Florax and Rey, 1995). Finally, re-specification of the weights matrix may eliminate the need for the SAC model, or any higher-order spatial model.

Goetz and Rupasingha do not use the LM test (see Appendix B), rather they use the SAC model and validate *ex-post* when they find rho and lambda are statistically significant. They write a lengthy interpretation of the rho and lambda coefficients, but this interpretation is invalid due to the identification problem discussed above.

Because Goetz and Rupasingha's spatial model specification is flawed, I follow Florax and Rey's (1995) LM procedure. This procedure identifies the spatial error process as dominant, thus the Spatial Error Model (SEM) is the appropriate spatial econometrics model. The SEM model is identical to the OLS specification, but I specify the non-spherical error term, ε, as:

(5.8) $\varepsilon = \lambda W \varepsilon + u$ where $u \sim i.i.d.$

5.2.2 Estimated Equations

I estimate a series of entrepreneurship models using SEM structure and a maximum likelihood estimator. I begin with the growth equation, which includes the lagged level of entrepreneurship (Equation 5.9), and I use *EI_se_chg, Prop_chg,* and *Prop_chg90s*, *Prop_chg* calculated as change between 1990 and 2000 as dependent variables.

$$
\begin{aligned}
& y_{t-(t-1)} = \beta_0 + \beta_1 HomeValue_{t-1} + \beta_2 HomeOwn_{t-1} + \beta_3 College_{t-1} \\
& + \beta_4 HS_{t-1} + \beta_5 MedAge_{t-1} + \beta_6 Female_{t-1} + \beta_7 White_{t-1} \\
& + \beta_8 WSinc + \beta_9 WSIncGro_{t-1} + \beta_{10} DeposPop_{t-1} + \beta_{11} Unemp_{t-1} \\
(5.9) \quad & + \beta_{12} Unemp^2{}_{t-1} + \beta_{13} Nonmetro \; + \beta_{14} Amenity \; + \beta_{15} Ag + \\
& + \beta_{16} Mining + \beta_{17} NonDurManu + \beta_{18} DurManu + \beta_{19} Trade + \\
& + \beta_{20} Visitor + \beta_{21} Services + \beta_{22} Tax_{t-1} + \beta_{23} Amenity + \beta_{24} y_{t-1} + \varepsilon,
\end{aligned}
$$

For the level dependent variable, I estimate the same model without the lagged-level (Equation 5.10). Dependent variables for the initial estimation include the *EI_se/emp, EI_birth/pop*, and *Prop*.

$$
\begin{aligned}
& y_t = \beta_0 + \beta_1 HomeValue_{t-1} + \beta_2 HomeOwn_{t-1} + \beta_3 College_{t-1} \\
& + \beta_4 HS_{t-1} + \beta_5 MedAge_{t-1} + \beta_6 Female_{t-1} + \beta_7 White_{t-1} \\
& + \beta_8 WSinc + \beta_9 WSIncGro_{t-1} + \beta_{10} DeposPop_{t-1} + \beta_{11} Unemp_{t-1} \\
(5.10) \quad & + \beta_{12} Unemp^2{}_{t-1} + \beta_{13} Nonmetro \; + \beta_{14} Amenity \; + \beta_{15} Ag + \\
& + \beta_{16} Mining + \beta_{17} NonDurManu + \beta_{18} DurManu + \beta_{19} Trade + \\
& + \beta_{20} Visitor + \beta_{21} Services + \beta_{22} Tax_{t-1} + \beta_{23} Amenity + \varepsilon,
\end{aligned}
$$

5.3 RESULTS AND SENSITIVITY ANALYSIS

5.3.1 Base Model: Growth in Entrepreneurship

I estimate the entrepreneurship growth model with Equation 5.9. Data for calculating growth in Entrepreneurial Industries applied to births are not available, so results are based on the estimation of change in Entrepreneurial Industries self employment between 2001-2006, *EI_se_chg*. A summary of results is presented in Table 5.3, and the full set of results is available in Appendix C.

Results suggest that natural amenities, access to financial collateral, and location in metropolitan statistical areas are the best predictors of growth in *EI_se_chg* in the model (Table 5.3, Model 1). The positive and significant (0.05) sign on *Amenity* affirms work by McGranahan and Wojan (2007) that argued amenities attract knowledgeable and skilled workers. I expected *Amenity* to have a positive relationship with Entrepreneurial Industries growth because these knowledgeable and skilled people are more likely to be innovative and entrepreneurial. Many studies have found that access to capital increases growth in entrepreneurship (Garofolli, 1994), and I do not find evidence to reject this hypothesis. Coefficient signs on home ownership and median home value were positive and significant (0.01), suggesting that where housing values are higher and more people owned a home, in 2000, Entrepreneurial Industries grew more. Correlations and spatial data analysis found Entrepreneurial Industries was highest in metropolitan counties, and the negative and significant coefficient on *Nonmetro* affirms the statistical significance of these findings.

Human capital, demographic, and seed capital variables behaved differently than expected. Prior work suggests that entrepreneurs are older, more likely to be male, educated, and Caucasian that the population as a whole. I find a negative coefficient on *MedAge*, a positive coefficient on *Female*, and a zero coefficient on *White*, suggesting that growth in Entrepreneurial Industries might be via non-traditional entrepreneurs who are younger, female, and less educated. These demographics are also characteristic of cities, so it is possible that *Nonmetro* did not control for these characteristics. Similarly, the negative coefficient on *College* may reflect that a larger percent of city residents are college educated, and Entrepreneurial Industries growth is highest outside these areas because the high school educated are necessary employees for the entrepreneur. Finally, the negative coefficient sign on *DeposPop* can be attributed to more financial sophistication in areas of high Entrepreneurial Industries growth—less local bank deposits could indicate more investments in the stock market, the business itself, or other, more sophisticated financial instruments. In conclusion, theoretically inconsistent coefficient signs signal the need for cautious interpretation of the results.

I found no relationship between unemployment and Entrepreneurial Industries growth. Although other studies have found a relationship between unemployment and entrepreneurship (Parker, 1996), my finding results from excluding necessity-based entrepreneurs from my entrepreneurship indicator.

Growth in wage and salary income during the 1990s had a positive coefficient while the level of wage and salary income had a negative coefficient. To explain the negative coefficient, I must assume that where wage and salary incomes were high there was less incentive to innovate or take the risks required to become an entrepreneur. The positive coefficient sign on wage and salary income growth during the 1990s suggests the region, as a whole, is experiencing economic growth and prosperity.

For a comparison, I run the same regression using Goetz and Rupasingha's dependent variable, growth in nonfarm proprietorships, *Prop_chg* (Table 5.3, Model 2), and I found a major difference in the determinants of *Prop_chg* and *EI_se/emp*. Indeed, the only significant coefficient that had the same sign as in Model 1 was *Nonmetro*, indicating that growth in both Entrepreneurial Industries and self employment was higher in metro counties than nonmetro counties. One difference of interest is the coefficient sign on *Amenity*; it had a negative coefficient whereas it is consistently positive and significant in the EI regressions, suggesting growth in self employment occurs in low-amenity areas. The adjusted R-square[11] for *Prop_chg* (model 2) is 0.042, but the adjusted R-square is four times as high, 0.173, for the *EI_se_chg* model (model 1), which suggests variables, such as financial collateral and amenities, may help drive innovative entrepreneurship.

Because my 2001-2006 results differ from Goetz and Rupasingha's, I run my model using *Prop_chg_90s* as the dependent variables and 1990 explanatory variables (Table 5.3, Model 3). Although Goetz and Rupasingha use the same dependent variable, my results are very different, likely because I corrected the spatial econometric model specification. Wage and salary income, and its growth all have a negative coefficient, suggesting growth in self employment during the 1990s was highest in areas that featured low wages and little or no growth in wages—all features of necessity-based entrepreneurship. Fit was higher for the 1990-2000 model proprietor growth model than the 2001-2006 model (adjusted R-square=0.099, 0.042, respectively), likely due to the tremendous growth over the 1990-2000 period. This difference shows that regression results can vary with the selected time period, which might explain some of the many discrepancies in entrepreneurship model results.

---

[11] I report the adjusted R-square of the OLS regression because the pseudo R-square from the Maximum Likelihood Estimation (MLE), correlation between response and fitted variables, is only a rough estimate of the explanatory power of the model, and does not have the same meaning that the R-square of a linear model has—making interpretation of the pseudo R-square difficult.

Table 5.3 Determinants of Growth in Entrepreneurial Industries and Entrepreneurship

	Model 1	Model 2	Model 3	
	EI_se_chg	Prop_chg	Prop_chg_90s	Expectation
HomeValue	+			+
HomeOwn	+			+
College	-			+
HS		-		+/-
MedAge	-	+		+
Female	+			-
White		-		+
WSinc	-		-	+/-
WSincGro	+		-	+/-
DeposPop	-			+
Unemp				+/-
UnempSq				+/-
Nonmetro	-	-	-	-
Ag			+	+
Mining				+/-
NonDurManu		+		+/-
DurManu		+		+/-
Trade				+/-
Visitor				+/-
Services		+		+
Amenity	+	-		+
Tax				+/-
Level of Y	-	+	-	+/-
*Adj. R^2**	*0.173*	*0.042*	*0.099*	

**OLS R-square, not pseudo R-square*

5.3.2 Determinants of the Entrepreneurial Industries Level

I model the level of entrepreneurship, *EI_se/emp* and *EI_birth/pop*, using the specification in Equation 5.10. By examining the results of Entrepreneurial Industries applied to both self employment and births, I hope to understand the drivers of both the stock of individual entrepreneurs and the flow of entrepreneurial establishments. A full set of results are in Appendix C.

In both models, *Amenity*, *College*, financial collateral, and growth in wage and salary income have a positive relationship with the level of Entrepreneurial Industries (Table 5.4). Like the growth model, *Amenity* has a positive and significant (0.01) relationship with Entrepreneurial Industries, suggesting that innovative entrepreneurs live and work in pleasant and/or scenic places. The percent of adults with a college education, *College*, is also positively related to Entrepreneurial Industries, but without data on individual entrepreneurs I cannot tell if the entrepreneurs themselves have college educations, or it the entrepreneurs live/operate near a skilled labor force. Percent of adults with a high school diploma was insignificant. Results suggest the availability of financial collateral, *HomeValue* and *HomeOwn*, and growth in wage and salary incomes, *WSincGro*, contribute to the level of Entrepreneurial Industries, suggesting that Entrepreneurial Industries is higher in socioeconomically advantaged counties. Unemployment variables were insignificant in both models, as expected, and observed in the Entrepreneurial Industries growth model.

The dummy variable for nonmetropolitan counties is insignificant in the *EI_birth/pop* model, although it is negative in the *EI_se/emp* model and the Entrepreneurial Industries growth models. I expected the coefficient sign to be negative because Entrepreneurial Industries was higher in metro counties. I do not find evidence of multicollinearity, which could lead to a wrong coefficient sign, so I conclude that, on the aggregate, Entrepreneurial Industries births are not significantly different in metro and nonmetro counties.

Table 5.4 Determinants of Entrepreneurial Industries and Entrepreneurship

	Model 4	Model 5	Model 6	
	EI_se/emp	EI_birth/pop	Prop	expectation
HomeValue	+	+	+	+
HomeOwn	+		+	+
College	+	+	+	+
HS				+/-
MedAge	-	+	+	+
Female	+		-	-
White	+		-	+
WSinc	-		-	+/-
WSincGro	+	+	-	+/-
DeposPop	-	+		+
Unemp			-	+/-
UnempSq				+/-
Nonmetro	-		-	-
Ag	-	+		+
Mining				+/-
NonDurManu			-	+/-
DurManu				+/-
Trade			-	+/-
Visitor		+	+	+/-
Services	-		-	+
Amenity	+	+		+
Tax		+		+/-
*Adj. R^2**	*0.357*	*0.398*	*0.511*	

**OLS R-square, not pseudo R-square*

The individual income tax index has a positive coefficient in the birth model, suggesting the lower the state income tax burden on individuals, the higher the level of Entrepreneurial Industries births. This coefficient was insignificant in the entrepreneurship growth model though, perhaps because the self employed are more interested in the tax advantages associated with small business ownership that they are about the additional tax burden.

Coefficient signs on demographic variables are mixed. Median age has a positive coefficient in the Entrepreneurial Industries birth model but a negative coefficient in the Entrepreneurial Industries self employment model, while *Female* and *White* are positive in the Entrepreneurial Industries self employment model but insignificant in the Entrepreneurial Industries birth model. The coefficient on *DeposPop* is also mixed.

Mixed signs do not tell us much about entrepreneurship as a whole, but I think they signal differences between the stock of self employed and the flow of establishment births. Mixed coefficient signs also suggest that modeling entrepreneurship does not always lead to definitive results, and my results should be interpreted as such.

For a comparison, I use *Prop* as a dependent variable in the same equation (see Table 5.4, model 6). Financial collateral and human capital variable coefficients are the same in both the Entrepreneurial Industries and *Prop* models, suggesting these positive relationships are robust to different entrepreneurship measures. Coefficients signs on other measures, however, including demographics, are opposite and do not tell us much about entrepreneurship. The coefficient on *Amenity* is insignificant. Finally, the relationship between proprietorships and income growth is negative, but positive for Entrepreneurial Industries and income growth; this finding suggests that self employment occurs in lower income counties, perhaps due to necessity rather than to bring innovation to the market.

5.3.3 Sensitivity of Results to Choice of Innovation Industries

Sensitivity of my results to the method used to select innovative industries is important because, given the differences between Entrepreneurial Industries and self employment results, I want to ensure my results are independent of the Entrepreneurial Industries method. Using ST2 and STP as dependent variables, I run Equation 5.10; all but one estimated coefficient are the same in sign and significance (Table 5.5). Coefficient signs for ST3 (model 4) and ST2 (model 7) models are identical and only *College* differs in the STP model (model 8). The coefficient on *College* is negative, likely due to the dominance of manufacturing industries in STP because manufacturing establishments generally need skilled laborers for assembly, but not necessarily a college educated labor force. Results summarized in Table 5.5 suggest regression results discussed in Section 5.3.2 are insensitive to the choice of specific industries. Full results are available in Appendix C.

Table 5.5 Comparison of Results From EI Methodologies

	Model 7	Model 8	Model 4
	ST2	STP	EI_se/emp (ST3)
HomeValue	+	+	+
HomeOwn	+	+	+
College	+	-	+
HS			
MedAge	-	-	-
Female	+	+	+
White	+	+	+
WSinc	-	-	-
WSincGro	+	+	+
DeposPop	-		-
Unemp			
UnempSq			
Nonmetro	-	-	-
Ag	-	-	-
Mining			
NonDurManu		+	
DurManu			
Trade			
Visitor	+		
Services	-	-	-
Amenity	+	+	+
Tax	+		
R-square	*0.542*	*0.346*	*0.357*

5.4 CONCLUSION

I found financial collateral, income growth, being in a metro area, and natural amenities drive regional entrepreneurship. These findings are a stark contrast to results using self employment to measure entrepreneurship. In short, regression results vary with the entrepreneurship metric used, leading to mixed and sometimes theoretically inconsistent results. Coefficient signs on demographic variables, in particular, had little or no consistency between models. Differences are likely due to the exclusion of necessity-based entrepreneurs. My findings illustrate the problems with using entrepreneurship model results to identify economic development and policy strategies. Mixed and theoretically inconsistent coefficient signs signal the need for cautious interpretation of the results.

Results also vary with the time period used. Tremendous growth in the 1990s likely created the differences in coefficient sign and fit between the 1991-2001 and 2001-2006 models of self employment. This difference raises concerns about the usefulness of entrepreneurship model results for creating policy recommendations and might explain some of the discrepancies among different models' results (Bruyat and Pierre-Andre, 2000; Tamasy, 2006).

Results suggest that different entrepreneurship metrics and time periods fuel policymaker confusion, making it difficult to discern valuable findings and questionable findings (Tamasy, 2006). Rather than continually tweaking models and metrics, I think future research on the determinants of entrepreneurship should be region-specific and policy recommendations based on regional strengths and weaknesses, using these regression results as only a starting point. I do not think we can learn much more from modeling entrepreneurship than we already have.

CHAPTER 6: ENTREPRENEURIAL INDUSTRIES: REGIONAL GROWTH MODEL

The widely held belief that entrepreneurship and long-term regional employment growth are correlated (Acs and Armington, 2003) has spurred a growing body of research examining the consequences of entrepreneurship on regional growth. An innovation-entrepreneurship-growth nexus is widely touted, yet it has not been established empirically (SBA, 2005). Omitting innovation from entrepreneurship measures has handicapped this growing body of research.

Recent research suggests entrepreneurship is a vehicle for incorporating human capital, research and development, and innovation into the economy (Acs et al., 2004; Glaeser, 2006). McGranahan, Wojan, and Lambert (2009) build on these ideas and examine how the nexus between entrepreneurship and creative class affects economic growth. They develop a model of county growth incorporating amenity levels and test to what extent the entrepreneurship and human capital drives nonmetropolitan growth in the presence of different amenity levels.

I use McGranahan et al.'s growth model as a starting point for my model because it is a parsimonious and state-of-the-art model that accounts for the nexus between amenities, skills, entrepreneurship, and growth. I proceed by discussing the model and the entrepreneurship metrics employed, which include Entrepreneurial Industries and McGranahan et al.'s entrepreneurship measures. I find Entrepreneurial Industries have a robust, positive relationship with economic growth and conclude that the best way of advancing entrepreneurship policy and practice is to use what we have already learned to start building region-specific solutions.

6.1. MODEL

McGranahan, Wojan, and Lambert's (2009) model differs from previous growth models by recognizing that knowledge and creativity are not intrinsic characteristics of places. They test whether the interaction between creative capital, a proxy for knowledge and talent, and entrepreneurship explains variation in nonmetro county growth, particularly in the context of different place-based amenity levels.

McGranahan et al. (2009) posit that outdoor amenities attract talent, but entrepreneurship is necessary to incorporate this talent (or set of skills and knowledge) into the economy to create growth in establishments, jobs, start-ups, and the creative class. McGranahan et al. find counties with a higher proportion of creative class *and* entrepreneurship experienced more growth during the 1990s than other counties. Results suggest the entrepreneurship/creative class nexus is particularly strong in high amenity areas, e.g., mountainous and coastal areas, but the relationship is less relevant in low amenity areas, e.g., the Great Plains.

McGranahan et al.'s model provides a solid foundation for my model because it is relatively parsimonious; authors found simultaneous estimation was unnecessary and this enables me to simplify the model and its interpretation so I can focus on the richness of my results. Finally, McGranahan et al.'s model incorporates two of the most popular measures of entrepreneurship—self employment and the establishment rate, so substituting-in Entrepreneurial Industries is a natural modification to the model.[12]

6.1.1. Explanatory Variables

I model growth as a function of P, local resources (including entrepreneurship), Λ, labor market characteristics, Υ, urban influence, Σ, industry sectors, Δ, demographic characteristics, I, institutions, and A, amenities (Equation 6.1). Table 6.1 contains variable definitions, sources, and summary statistics. All explanatory variables are for the year 2000, unless otherwise noted.

(6.1) $$Growth_{06-01} = \mathrm{P} + \Lambda + \Upsilon + \Sigma + \Delta + \mathrm{I} + \mathrm{A} + \varepsilon,$$

6.1.1.1 Local resources vector, P, and test variables

Rho, P, represents the vector of local resources and includes the test variable, entrepreneurship (Table 6.1, denoted in grey). Entrepreneurial Industries variables are

[12] McGranahan et al test, independently, two measures of entrepreneurship—self employment and the establishment rate. *Self employment* is one of the most widely used measures of entrepreneurship but it overestimates entrepreneurship because it does not capture the innovative component of entrepreneurship. The establishment rate is the ratio of establishments to employees, the inverse of the widely used *average employee per establishment* measure of entrepreneurship. This measure is problematic because it does not capture innovation, risk, and uncertainty.

EI_se/emp and *EI_birth/pop*, and both are discussed in detail in Chapter 4. Both are better indicators of entrepreneurship than widely used metrics because they capture innovation, a key component of entrepreneurship (Schumpeter and Opie, 1983).

Other variables in the local resources vector include percent recast creative class, *Creative,* as discussed in McGranahan and Wojan (2007), percent of adults over 25 with a high school diploma, *HS*, and the percent of adults over 25 with a four-year college degree, *College,* to control for the level of human capital. *Creative* and the entrepreneurship variable are standardized to aid interpretation, and I expect them to have positive coefficients (McGranahan et al., 2009). Finally, the interaction between *Creative* and the entrepreneurship measure is included because it is McGranahan et al.'s test variable.

Lambda, Λ, represents the vector of labor market explanatory variables and includes the employment rate, *EmpRate*, and median household income, *MedInc*. McGranahan et al. (2009) used the employment rate rather than the unemployment rate, arguing that underemployment and discouraged workers often skew the unemployment rate downward in rural areas. Although discouraged workers also affect the employment rate, McGranahan et al. argue that it is less affected by them.

Upsilon, Υ, represents the vector of urban influence variables and includes population density, *PopDen*, the percent of workers working outside the county, *Commute*, and a dummy variable for metropolitan counties, *Metro*. This vector is included because previous research indicated that growth is higher in densely settled areas, likely due to larger labor pools (McGranahan and Wojan, 2007).

Sigma, Σ, represents the vector of industry employment variables. Industry employment is calculated as the percent of employed persons employed in each industry. The model includes *Ag*, agriculture and forestry, *Mining*, *NonDurManu*, non-durables manufacturing, *DurManu*, durables manufacturing, *Trade*, wholesale and retail trade, *Visitor*, recreation, arts, accommodation, and food services, and *Services*, information, finance, insurance, real estate, and professional, scientific, and technical services. Other industries are the omitted condition. *InnovEmp*, the percent of employment in innovative industry establishments, controls for the presence of innovative industries in each county, and ensures my test variable only captures the nexus between innovative industries and entrepreneurs.

Delta, Δ, represents the vector of demographic variables. The population aged 8-17, *Pop8-17*, represents the future labor force, population over the age of 62, *Pop62*, controls for areas that attract many retirees, and percent black, *PctBlack*, percent Native American, *PctNA*, and percent Hispanic, *PctHis* are included because different groups may have different opportunities and proclivities to engage in economic activity (McGranahan et al., 2009).

Iota, I, represents the vector of institutional variables, which control for employment affects due to the presence of large institutions. Institutional variables include *Military*, percent aged 18-24 who are serving in the armed services, and the percent of the population aged 18-62 who are currently enrolled in higher education, *CollegePop*.

Alpha, A, represents the vector of amenity variables, outdoor amenities, *OutAmen*, and public land, *PubLand*. The outdoor amenities variable is similar to the widely used amenity variable (McGranahan, 1999), but it includes landscape—percent forest and its square—because recent literature indicates landscape preference for partially forested areas (McGranahan, 2008). For details on how *OutAmen* is constructed, see McGranahan et al. (2009). *PubLand* is the percent of land in each county publically owned, based on a survey by the U.S. Forest Service.

Finally, I include the lag of the three dependent variables, change in establishments, *EstabChg90,* change in jobs (nonfarm employment), *JobChg90*, and change in population, *PopChg90*. State fixed effects are included, and Alabama is the omitted condition.

Table 6.1 Explanatory Variables for Growth Model

Explanatory Variables			Source	Mean	StD	Min	Max
P	*Entrepreneurship Test Variables (all are standardized)*						
	SelfEmp	nonfarm proprietor employment/total nonfarm employment	*	-1.9E-10	1	-3.8171	6.809
	Estab/Emp	private nonfarm establishments/private nonfarm employment	^	8.9E-10	1	-2.4467	5.110
	EI_se/emp	EI applied to self employment data" /nonfarm employment*	see text	6.6E-09	1	-0.8237	5.822
	EI_birth/pop	EI applied to births% /population*, 3-year MA 1999-2001	see text	3.1E-09	1	-1.1490	10.008
	ST2	ST2 industries applied to self emp data" /nonfarm emp*	see text	4.8E-09	1	-0.9415	6.302
	STP	STP industries applied to self emp data" /nonfarm emp*	see text	7.2E-09	1	-0.8142	5.873
	Creative	Creative class employment /total employment#, standardized	see text	3.6E-10	1.000	-2.922	6.345
	HS	% of population >age 25 with secondary school diploma/GED	#	0.774	0.087	0.347	0.970
	College	% of population over age 25 with a 4 year college degree	#	0.163	0.075	0.049	0.605
Λ	*EmpRate*	% of population age 16-64 employed	#	0.708	0.093	0.215	0.935
	MedInc	Median household income	#	35021	8604	9888	82929
Υ	*PopDen*	Population/land area	#	214	1520	0.0966	54235
	Commute	% of employed working out of county	#	0.320	0.173	0.017	0.862
	Metro	OMB-designated metropolitan county, 2003	OMB	0.338	0.473	0	1
Σ	*Ag*	% of employed in agriculture & forestry	#	0.063	0.071	0	0.556
	Mining	% in mining	#	0.012	0.027	0	0.456
	NonDurManu	% in nondurable manufacturing	#	0.071	0.051	0	0.431
	DurManu	% in durable manufacturing	#	0.089	0.064	0	0.420
	Trade	% in retail and wholesale trade	#	0.145	0.025	0.017	0.299
	Visitor	% in arts, recreation, food, and accommodation	#	0.071	0.033	0	0.364
	Services	% in information, FIRE, Prof & Sci Services	#	0.106	0.044	0	0.411
	InnvInd_Emp	% emp. in Innovative Industry establishments, 2000, stdzd	^	0.019	0.031	0	0.253
Δ	*Pop8-17*	% of population age 8-17	#	0.181	0.021	0.097	0.308
	Pop62	% of population age 62 and over	#	0.175	0.046	0.024	0.397
	PctBlack	Black % of population	#	0.083	0.143	0	0.861
	PctNA	Native American % of population	#	0.016	0.064	0	0.937
	PctHis	Hispanic % of population	#	0.063	0.123	0	0.981
I	*Military*	% of population 18-24 in the Armed Services	#	0.004	0.024	0	0.610
	CollegePop	% of population 18-64 enrolled in college or university	#	0.079	0.057	0.010	0.539
A	*OutAmen*	climate and landscape measure	see text	-0.054	0.953	-2.094	4.696
	PubLand	Public % of land area, stdzd, US Forest Service	see text	-0.036	0.943	-0.582	5.401
Y_	*EstabChg90*	Log change in establishments, 1990-2000	^	0.016	0.033	-0.223	0.546
	JobChg90	Log change in employment, 1990-2000	*	4.788	0.159	4.190	6.761
	PopChg90	Log change in in population, 1990-2000	#	0.0013	0.019	0	0.910

All data are 2000 unless otherwise noted

* Bureau of Economic Analysis, Regional Economic Information System (REIS)

^ Bureau of Census, County Business Patterns

" Bureau of Census, Nonemployer Statistics, 2006, unless otherwise noted

\# Decennial Census of Population, 2000, unless otherwise noted

% Special tabulation of single unit employer establishment births

6.1.2 Dependent Variables

Like McGranahan et al. (2009), I use change in jobs and change in establishments as dependent variables. I also use change in population because it has been widely used as a dependent variable in growth studies. Following McGranahan et al., I calculate the dependent variables, a growth rate, as *t+1* minus t, normalized by *t*. I calculate growth between 2001, *t*, and 2006, *t+1*, to proxy for the most recent, 2001-2007, business cycle, because 2007 are unavailable. Change in population, *PopChg*, is calculated using BEA-REIS data. Change in employment, or jobs, *JobChg*, is calculated using BEA-REIS also, and change in establishments, *EstabChg*, with *County Business Patterns* data. Table 6.2

contains the variable names, descriptions, source, and summary statistics for the three dependent variables used in this analysis.

Table 6.2 Dependent Variables

Dependent Variables			*Mean*	*StDev*	*Min*	*Max*
PopChg	Change in population, 2001-2006/population, 2001	*	0.0149	0.0755	-0.792	0.535
EstabChg	Change in establishments, 2001-2006/private sector nonfarm jobs, 2001	^	0.0245	0.0885	-1	0.537
JobChg	Change in nonfarm jobs, 2001-2006/nonfarm jobs 2001	*	0.0767	0.0963	-0.358	1.116

* BEA-REIS

^ U.S. Census Bureau, County Business Patterns

6.1.3 Model Specification

I specify the OLS model using the variables discussed above and use the results to test for multicollinearity, heteroskedasticity, and spatial dependence. I do not find evidence of multicollinearity among the explanatory variables. I do find, however, evidence of heteroskedasticity in the OLS estimation, as indicated by the Breusch-Pagan test ($BP=1811.5$, and $p<0.001$). Heteroskedasticity is very common in U.S. county-level regressions due to the heterogeneity among counties. I use the White-Huber correction to make the standard errors robust to heteroskedasticity and find that the recalculated t-statistics on the Entrepreneurial Industries indicators are smaller, although they all remain statistically different from zero.

Increasingly, regional growth models control for spatial effects because growth processes vary widely across the United States and can cause coefficients to be misinterpreted (Partridge et al., 2008). The dynamics of rural and urban growth vary, and county heterogeneity makes the problem especially complex (Feser and Isserman, 2006; Partridge et al., 2008). Administrative boundaries, the degree of agglomeration, and rural-urban interaction affect the direction and magnitude of growth and change.

The presence of heteroskedasticity suggests a spatial dependence problem in the OLS residuals, so I conduct Lagrange Multiplier (LM) tests for spatial dependency structure using the procedure in Appendix B.4. I use a first-order queen contiguity matrix in the tests due to the nature of spatial dependence and its suitability for use with irregular polygons. LM and Robust LM tests indicate the Spatial Error Model (SEM) is appropriate; it will control for nuisance errors, which override substantive error processes, and reduce heterogeneity.

I estimate Equation 6.2 with the explanatory variables presented in Table 6.1 and I specify the structure of epsilon to be consistent with SEM. Let *ESHIP* represent any one of the entrepreneurship variables described in Table 6.1, *EshipXcc*, the interaction term between the *ESHIP* and *Creative,* and *Growth*, any of the three dependent variable presented in Table 6.2.

$$
\begin{aligned}
(6.2)\quad & Growth_{06-01} = ESHIP_{00} + Creative_{00} + EshipXcc_{00} + College_{00} + HS_{00} \\
& + EmpRate_{00} + MedInc_{00} + PopDen_{00} + Commute_{00} + Metro + Ag_{00} \\
& + Mining_{00} + NonDurManu_{00} + DurManu_{00} + Trade_{00} + Visitor_{00} \\
& + Services_{00} + Age8-17_{00} + Age62_{00} + PctBlack_{00} + PctNA_{00} + PctHis_{00} \\
& + MilitaryPop_{00} + CollegePop_{00} + OutAmen + PubLand + EstabChg90_{00-90} \\
& + JobChg90_{00-90} + PopChg90_{00-90} + StateFE + \varepsilon,
\end{aligned}
$$

where $\varepsilon = \lambda \mathbf{W} \varepsilon + \mu$ and μ is assumed independently and identically distributed (i.i.d.).

6.2 RESULTS AND SENSITIVITY ANALYSIS

6.2.1 Estimation Results

Entrepreneurial Industries has a positive relationship with population, employment, and establishment growth. Results are summarized in Table 6.3 with dependent variables on the left-hand-side. Standardized coefficients on Entrepreneurial Industries are relatively close to each other, but lowest for the *EstabChg* regressions, perhaps because establishment formation is an employment strategy in weaker economies, like self employment by necessity (McGranahan et al., 2009).

Table 6.3 Summary of Growth Model Results Using Entrepreneurial Industries

	Entrepreneurship Variable							
	EI_se/emp				EI_birth/pop			
Dependent Variable	Coef	Z		R^2	Coef	Z		R^2
Δpop	0.017	12.4	***	0.58	0.004	3.2	***	0.57
Δjobs	0.013	6.2	***	0.38	0.002	1.0		0.39
Δestabs	0.010	5.0	***	0.40	0.005	2.5	***	0.4

The only insignificant coefficient between growth and Entrepreneurial Industries is the coefficient on *EI_birth/pop* with dependent variable, *JobChg*, which is odd because theoretically, the birth of an establishment necessitates at least one paid employee. Although the coefficient on *EI_birth/pop* is positive, it is not statistically different from zero. A plausible explanation is that a nascent single-unit establishment has one paid employee but two unpaid proprietors; if the unpaid proprietors have to drop out of the wage and salary job market to start the business, then the birth is accompanied by the loss of two jobs, and on the aggregate, an insignificant number of jobs are created. This hypothesis might be especially true during the early 2000s due to the jobless recovery, but without more data on business cycle effects and flow data fluctuation I cannot test this explanation.

I add percent employment in innovative industries (*InnovEmp*) to the model to test if innovative industries, not the industry/entrepreneurship nexus, drive results and find no evidence to support this. Entrepreneurial Industries remains positive and significant (0.01) and coefficient size does not decrease when I add *InnovEmp* to the model, suggesting that the innovative industries and self employment/birth nexus is a unique driver of growth.

For a comparison, I test the relationship between growth and traditional entrepreneurship measures, *SelfEmp* and *Estab/Emp,* in the same model; results were different and inconsistent with theory (Table 6.4). Although *SelfEmp* and *Estab/Emp* both have a positive relationship with *JobChg,* neither have a statistically significant relationship with *EstabChg*. The most troubling results is that *PopChg* has a negative relationship with the self employment rate and no relationship with the establishment rate, perhaps because self employment is highest in sparsely populated areas and *SelfEmp* includes entrepreneurship by necessity. McGranahan et al. did not use change in population as a dependent variable, so I cannot compare these unexpected results to theirs. My findings illustrate the problems with modeling the relationship between growth and entrepreneurship, the results vary with the chosen entrepreneurship measure and time period selected.

Table 6.4 Summary of Growth Model Results Using Widely Used Entrepreneurship Measures

		Entrepreneurship Variable							
		SelfEmp Rate				Estab/Emp			
		Coef	Z		R^2	Coef	Z		R^2
Dependent Variable	Δpop	-0.007	-5.3	***	0.56	-0.001	-0.8		0.56
	Δjobs	0.019	9.2	***	0.40	0.015	7.2	***	0.39
	Δestabs	-0.003	-1.5		0.39	-0.001	-0.8		0.40

Significance Level: *** 0.01 ** 0.05 * 0.1

6.2.2 Sensitivity of Results to Methodology

I find the positive relationship between Entrepreneurial Industries and growth is robust to variation in the Entrepreneurial Industries method (Table 6.5). The coefficients on ST2 (skill and technology at two times the mean, secondary criteria the same as ST3) and STP (skill, tech, and patents are the primary criteria, no secondary criteria) are positive, and slightly larger than ST3 (skill and technology at three times the mean), possibly because both include more establishments than ST3. AIC scores are lowest for the alternative measures. The adjusted R-square values are highest for the base model, ST3, which suggests the model captures more of the variation in ST3 than other indicators and indicates ST3 results may be of more use for policy and practice. Full model results are available in Appendix D.

Table 6.5 Sensitivity to Entrepreneurial Industries Methodology

Dependent Variable	ST2 Methodology Coef	Z		R^2	AIC	STP Methodology Coef	Z		R^2	AIC
Δpop	0.021	12.4	***	0.49	-9063	0.018	12.4	***	0.51	-9031
Δjobs	0.019	7.4	***	0.29	-6454	0.014	6.1	***	0.29	-6435
Δestabs	0.016	6.6	***	0.29	-6981	0.010	5.0	***	0.28	-6953

Dependent Variable	Base Eqn-ST3 Coef	Z		R^2	AIC
Δpop	0.017	12.4	***	0.58	-9027
Δjobs	0.013	6.2	***	0.38	-6435
Δestabs	0.010	5.0	***	0.40	-6951

6.3 CONCLUSION

Entrepreneurial Industries has a robust positive relationship with growth in population, employment, and establishments. These findings are consistent with my expectations, likely because Entrepreneurial Industries captures the innovative nature of entrepreneurship, which others have found is associated with economic growth (SBA, 2005). The choice of entrepreneurship measure can affect results. Some measures do not lead to theoretically consistent results. I find a negative relationship between *PopChg* and the self employment and establishment rate, both of which are widely used entrepreneurship measures. I do not find a statistically significant relationship between *EstabChg* and the widely used entrepreneurship measures. Widespread use of such non-innovative measures may be causing policymaker confusion (Tamasy, 2006). My findings suggest that Entrepreneurial Industries is a better indicator of entrepreneurship because it produced theoretically consistent results.

Before researchers spend more time and effort fixing specification problems, endogeneity problems, and the spatial econometric specification, I encourage them to take a step back and look at the big-picture, regional growth modeling. I have shown that the relationship between entrepreneurship and growth changes with both the definition of entrepreneurship and with the definition of growth. These results beg the questions, "what can we learn from these exercises" and "how useful is this growth model to policymakers and economic development practitioners."

Future research on regional economic growth should be region-specific. Results, such as these, can be used as a starting-point, but the uniqueness of each region suggests that a one-size-fits-all recipe for economic growth is a dream. Instead of continually striving to improve, or tinker with, the study of entrepreneurship and its effect on economic growth, researchers should consider focusing on region-specific work and interpreting results we already have for use in different regions.

CHAPTER 7: CONCLUSION

The answers to our problems don't lie beyond our reach.
They exist in our laboratories and universities; in our fields and our factories;
in the imaginations of our entrepreneurs.
—President Barack Obama, Inauguration Day, January 20, 2009

I have created an indicator of entrepreneurship that captures multiple attributes of entrepreneurship, including innovation—an aspect of entrepreneurship that contributes to economic growth, but is ignored by existing measures. *Entrepreneurial Industries* is a valid indicator of entrepreneurship. Entrepreneurial Industries is a refinement of widely utilized entrepreneurship measures and is available annually at the county-level. Entrepreneurial Industries also better represents Schumpeterian entrepreneurship and captures multiple dimensions of entrepreneurship (stock/flow, individual/establishment) with two metrics. Entrepreneurial Industries has the potential to improve regional entrepreneurship research by enabling it to focus on innovative entrepreneurship. Entrepreneurial Industries also has an audience in economic development practitioners and policymakers who strive for recent, relevant data and benchmarks to better guide policymaking.

This dissertation makes broader contributions to regional economic research. I discuss measure standardization, raise questions about the robustness of other, widely-used measures, discuss the effects of the selected time period, and develop methods to identify high skill occupations and innovative industries. Moreover, budding spatial econometricians can use the appendix on spatial econometrics and learn from the higher-order model discussion. This chapter reviews the merits of Entrepreneurial Industries and other contributions of the dissertation.

7.1 ENTREPRENEURIAL INDUSTRIES: WHAT IT IS

The main contribution of Entrepreneurial Industries is that it is a valid indicator of innovation—Schumpeter's concept of entrepreneurship—while other widely available county-level measures do not consider entrepreneurship. The nexus of innovative industries and births/self employment differs from total innovative industry employment and parent entrepreneurship measures so Entrepreneurial Industries is a better indicator of entrepreneurship for policy and economic development work because it offers a useful benchmark indicator. For FY 2010, U.S. States have budgeted $42.3 million for entrepreneurship development programs (C2ER, 2009).[13] Improving the entrepreneurship benchmark for the programs and their supporting policies could lead to more effective economic development and increase economic growth at minimal cost. An entrepreneurship indicator that captures innovation is also useful for regional researchers who have long noted the need for such an indicator.

Scholars posit that the ideal indicator of entrepreneurship must include multiple dimensions (Audretsch, 2005) and unlike other entrepreneurship indicators, Entrepreneurial Industries is multi-dimensional. The definition of entrepreneurship I posit in Chapter 2 serves as its foundation and requires that Entrepreneurial Industries meet three attributes of entrepreneurship, whereas most measures only capture one or two. Additionally, I can assess the stock of individuals and the flow of establishments with Entrepreneurial Industries, which is beneficial because stock and flow and individual and establishment measures can vary significantly.

Entrepreneurial Industries uses data available annually for counties. This makes it flexible and current enough for use by policymakers, economic development practitioners, and researchers. Counties are a good unit of analysis because they are at the heart of local policy and can be aggregated to labor market areas or metro areas. Annual availability makes the Entrepreneurial Industries indicator more flexible and timely than decennial Census data. Annual data are useful for policy applications that require recent data. The indicator is NAICS-based, so not available prior to 1997, limiting its potential for longitudinal analyses.

Entrepreneurial Industries has a robust, positive relationship with economic growth. Using the Entrepreneurial Industries indicator in an entrepreneurship model can help identify incentives and policy-levers for encouraging economic growth.

---

[13] Council for Community and Economic Research, *State Economic Development Database,* "Total State Expenditures by Functional Economic Development Program Area for FY2010". Thirteen States have clearly identified "Entrepreneurial Support" programs. This figure does not include federal funding or state funding t may be funneled through regional development agencies.

7.2 ENTREPRENEURIAL INDUSTRIES: WHAT IT IS NOT

Entrepreneurial Industries does possess several shortcomings. Self employment data are suppressed to prevent disclosure of individuals, a potential problem with EI_*se/emp* in sparsely populated counties. Accessing unsuppressed data requires a special agreement with the Census Bureau, like the one executed for the birth data. Another potential problem is the large year-to-year variation in flow data (birth data), which makes longitudinal data sensitive to individual observations. I use a moving-average for births that helps to avoid this problem, curtailing false-positives. Both of these issues relate to the indicator's use in rural areas; regression results for rural areas are the same despite these problems, but the coefficients are smaller. Where suitable, aggregating counties up to labor market areas may provide answers to these questions and eliminate these rurality-based problems.

This dissertation is motivated, in part, to reduce the confusion surrounding entrepreneurship measures. That said the complexity of the Entrepreneurial Industries indicator might *add* to policymaker and economic development practitioner confusion.

7.3 FUTURE WORK

Additional data could strengthen Entrepreneurial Industries, but the availability of these data is outside my immediate locus of control. Unsuppressed Nonemployer Statistics would bolster the self employment measure, particularly in rural areas. A longer time-series would improve evaluation of the measures. Micro data on innovative industries would enable me to examine the employment, value added, and productivity of these industries as well as the education, occupation, and industry of employment for individuals. Longitudinal information about innovative industry establishments and their exports, growth, and resiliency to downturn would also provide useful information on creating economic growth via entrepreneurship. I am currently working with the Center for Economic Studies in Suitland, MD, to gain access to some of these data for a firm resiliency project.

7.4 CONCLUSION

Entrepreneurial Industries is a valid indicator of the entrepreneurship construct. I hope this research encourages discussion and research on entrepreneurship measures and measure construction and clarity. It is timely and relevant, and Entrepreneurial Industries may be a good starting point to reinvigorate the literature.

While working on this dissertation, I learned to question data, question its use, and be suspicious when authors do not discuss method or measure construction. I found that the period for analysis could affect results, e.g., self employment during the 1990s and self employment during the 2000s. Measure construction, even the method used to standardize a rate, can affect results. "Results" can be shaped by the data and methods utilized. I learned that results can vary based on the metrics and model and, as a result, researchers should be wary of making policy recommendations based on one set of results. I hope researchers and end-users question available data and use the best possible entrepreneurship indicator because of this dissertation. For example, the differences between regression results using Entrepreneurial Industries and its parent measures are astounding.

Beyond discussion of the appropriate entrepreneurship metrics, I think entrepreneurship research is at a crossroad and needs to head in a different direction. We cannot learn much more from entrepreneurship and growth models because I think they are too sensitive to choice of data and model specification. For years, we have been tweaking econometric models and available data, however contradictory and theoretically inconsistent the end-product. Policymakers and practitioners are presented with confounding results, which end up being ignored.

Research needs to move away from tweaking models and focus on region-specific strategies based on places' strengths and weaknesses. Region-based entrepreneurship and economic development policies must be based on underlying regional research. Results of the entrepreneurship and growth model presented in this dissertation can be used as a starting-point, but the uniqueness of each region suggests that no "cookbook" answer to economic development problems exists.

REFERENCES

Acs, Zoltan, and Catherine Armington. 2003. "Endogenous Growth and Entrepreneurial Activities in Cities," Working Paper 03-02, U.S. Bureau of the Census, Center for Economic Studies.

Acs, Zoltan and Catherine Armington. 2006. *Entrepreneurship, Geography, and American Economic Growth.* New York: Cambridge University Press.

Acs, Zoltan, Catherine Armington and Ting Zhang. 2006. "The Determinants of New-firm Survival Across Regional Economies," Working Paper No. 0407, Max Plank Institute of Economics.

Acs, Zoltan J. and Pamela Mueller. 2007. "Employment Effects of Business Dynamics: Mice, Gazelles and Elephants," *Small Business Economics,* 30, 85-100.

Acs, Zoltan, Edward Glaeser, Robert Litan, Lee Fleming, Stephan Goetz, William Kerr, Steven Kelpper, Stuart Rosenthal, Olav Sorenson, and William Strange. 2008. "Entrepreneurship and Urban Success: Toward a Policy Consensus," Ewing Marion Kauffman Foundation. (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1092493).

Acs, Zoltan, William Parsons, and Spencer Tracy. 2008. "High-Impact Firms: Gazelles Revisited." Summary Report No. 328, *Small Business Research.*

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models.* The Netherlands: Kluwer Academic Publishers.

Anselin, Luc. 2008a. "Spatial Regression," in A. Stewart Fotheringham and Peter Rogerson (eds.), *Handbook of Spatial Analysis.* London: Sage, Ch. 26.

Anselin, Luc. 2008b. "*Model Specification,*" Lecture delivered at 21st Annual European Regional Science Association Summer School, Pécs, Hungary.

Audretsch, David B. 2002. "Entrepreneurship: A Survey of the Literature." Brussels: European Commission, Enterprise Directorate General.

Audretsch, David B. and Michael Fritsch. 1994. "On the measurement of entry rates," *Empirica,* Vol 21(1), 105-113.

Audretsch, David B. and Michael Fritsch. 1995. "The Measurement of Entry Rates: Reply" *Empirica,* 22, 159-161.

Audretsch, David B. and Michael Fritsch. 2002. "Growth Regimes over Time and Space," *Regional Studies,* 36, 113-124.

Audretsch, D., M.A. Carree, A.J. van Stel and A.R. Thurik. 2002. "Impeded Industrial Restructuring: The Growth Penalty," *Kyklos*, 55, 81-98.

Bartik, Timothy J. 1991. *Who Benefits From State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Baumol. William. 1990. "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy,* 98, pp. 893–921

Baumol, William. 1993. "Formal Entrepreneurship Theory in Economics: Existence and Bounds," *Journal of Business Venturing,* 8, 197-210.

Bednarzik, Robert W. 2000. "The role of entrepreneurship in U.S. and European job growth," *Monthly Labor Review*, 123, 3-16.

Birch, David L. 1981. "Who Creates Jobs?" *The Public Interest,* 65, 3-14.

Blanchflower, David. 2000. "Self-Employment in OECD Countries," *Labour Economics,* 7, 471-505.

Blanchflower, David. 2004. "Self Employment: More May Not Be Better," Working Paper No. 10286, National Bureau of Economic Research.

Blanchflower, David, and Andrew Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics*, 1, 26-60.

Bruyat, Christian, and Julien, Pierre-Andre. 2000. "Defining the Field of Research in Entrepreneurship," *Journal of Business Venturing,* 16, 165-180.

Bygrave, William D., and Charles Hofer. 1991. "Theorizing about Entrepreneurship," *Entrepreneurship Theory and Practice,* 16, 3-22.

Bull, I. and G.E. Willard. 1993. "Toward a Theory of Entrepreneurship," *Journal of Business Venturing,* 8, 183-195.

BEA. 2006. Bureau of Economic Analysis. Regional Economic Information System, 1969-2006. Available at: http://www.bea.gov/regional/docs/reis2006dvd.cfm

Cantillon, R. 1964 [1755] *Essai Sur La Nature Du Commerce En General.* (Translated by R. Higgs.) New York: A.M. Kelley Publishers, 1964 [1755].

Carlino, Gerald A., Chatterjee, Satyajit and Robert M. Hunt. 2007. "Urban Density and the Rate of Invention," *Journal of Urban Economics,* 61, 389-419.

Case, John. 1990. "The Most Entrepreneurial Cities in America," *Inc. Magazine,* March 1990. (http://www.inc.com/magazine/19900301/5071.html)

Casson, Mark. 2003. "The Entrepreneur: an economic theory," Cheltenham, UK: Edward Elgar.

Center for Regional Economic Competitiveness. 2009. "Siouxland Labor Force, Education and Training Assets," Prepared for Tri-State Regional Innovation Grant.

Chandra, Siddharth. 2002. "A Test of the Regional Growth-Instability Frontier Using State Data." *Land Economics*, 78:3, 442-462.

Chinitz, B. 1961. "Contrasts in Agglomeration: New York and Pittsburgh." *American Economic Review,* 51, 279-289.

Coase, R. H. 1937. "The Nature of the Firm," *Economic,* 4, 386-405.

Cole, Arthur H. 1942. "Entrepreneurship as an Area of Research," *The Journal of Economic History,* 2, 118-126.

C2ER. 2009. *State Economic Development Database.* Arlington, VA: Council for Community and Economic Research (C2ER.org).

de Wit G. 1993. "Models of self-employment in a competitive market," *Journal of Economic Surveys*, 7, 367-397.

Eff, Anthony. 2007. "Veblen in the Metropolis: Land Use Proximity in United States Urban Landscapes," paper presented at the North American Meetings of the Regional Science Association International, Savannah, GA.

Entrepreneur Magazine. 2006. "Hot Cities 2006", (www.entrepreneur.com/worklife/article165714-1.html)

Evans, David S. and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy,* 97, 808-827.

Evans, David S. and Linda S. Leighton. 1989. "The Determinants of Changes in U.S. Self-Employment, 1968-1987," *Small Business Economics,* 1, 111-119.

Fairlie, Robert W. 2009. "Kauffman Index of Entrepreneurial Activity 1996-2008," report prepared for Ewing Marion Kauffman Foundation.

Feser, Edward J. 2003. What Regions Do Rather than Make: A Proposed Set of Knowledge-based Occupation Clusters. *Urban Studies. Vol.40, No. 10, 1937-1958*

Feser, Edward and Andrew Isserman. 2006. "Harnessing Growth Spillovers for Rural Development: The Effects of Regional Spatial Structure," report prepared for the Office of the Under Secretary for Rural Development, U.S. Department of Agriculture, under Cooperative Agreement AG RBCS RBS-02-12.

Florax, R.J.G.M. and S. Rey. 1995. "The impact of misspecied spatial interaction structure in regression models," in L. Anselin and R.J.G.M. Florax (eds.) *New Directions in Spatial Analysis*, Berlin: Springer Verlag.

Garofolli, Gioacchino. 1994. "New Firm Formation and Regional Development: The Italian Case." *Regional Studies,* 28:4, 381-393.

Gartner, William B. 1990. "What are we talking about when we talk about entrepreneurship," *Journal of Business Venturing,* 5, 15-28.

Gartner, William B. and Scott A. Shane. 1995. "Measuring Entrepreneurship Over Time". *Journal of Business Venturing*, 10, 283-301.

Gebremariam, Gebremeskel, Gebremedhin, Tesfa and Peter V. Schaeffer. 2006. "An Empirical Analysis of County-Level Determinants of Small Business Growth and Poverty in Appalachia: A Spatial Simultaneous-Equations Approach," Working Paper No. 2006-3, West Virginia University, Regional Research Institute.

Georgellis, Yannis, and Howard J. Wall. 2000. "What makes a region entrepreneurial? Evidence from Britain," *The Annals of Regional Science,* 34, 385-403.

Glaeser, Edward. 2007. "Entrepreneurship in the City." Working Paper No. 13551, National Bureau for Economic Research.

Goetz, Stephan J. and David Freshwater. 2001. "State-level Determinants of Entrepreneurship and a Preliminary Measure of Entrepreneurial Climate," *Economic Development Quarterly,* 15, 58-70.

Goetz, Stephan and Anil Rupasingha. 2008. "Determinants of growth in non-farm proprietor densities in the U.S., 1990-2000," *Small Business Economics*, 32, 425-438.

Goetz, Stephan J. and Sundar S. Shrestha. 2009. "Explaining Self-Employment Success and Failure: Wal-Mart vs. Starbucks, or Schumpeter vs. Putnam," *Social Sciences Quarterly,* 91, 22-38.

Green, Gary Paul, Greg Wise, and Evan Armstrong. 2007. "Inventor and Entrepreneur Clubs: Investment in an Innovative Approach to Entrepreneurship," Paper presented at "Frameworks for Entrepreneurship Research in Food, Agriculture, Natural Resources and Rural Development: A National Conference on Entrepreneurship Research," Kansas City, Missouri.

Hamilton, Barton H. 2000. "Does Entrepreneurship Pay? An empirical analysis of the returns to self employment," *Journal of Political Economy.* 108, 604-631.

Hecker, Daniel. 2005. "High-technology employment: a NAICS-based update," *Monthly Labor Review*, 128, 57-72.

Henderson, Jason, Low, Sarah, and Stephan Weiler. 2007. "The Drivers of Regional Entrepreneurship in Rural and Metro Areas." In a Norman Walzer (ed.) *Entrepreneurship and Local Economic Development.* Ladham, MD: Lexington Books, 81-102.

Hoffmann, A., Larsen, M. and Oxholm, S. 2006. Quality Assessment of Entrepreneurship Indicators, FORA, Copenhagen. Available at: http://ice.foranet.dk/upload/quality_assessment_of_entrepreneurship_indicators002.pdf

Isserman, Andrew M. 2006. "In the National Interest: Defining Rural and Urban Correctly in Research and Public Policy," *International Regional Science Review,* 28, 465-499.

Iversen, Jens, Jorgensen, Rasmus and Mikolaj Malchow-Moller. 2008. "Defining and Measuring Entrepreneurship," *Foundations and Trends in Entrepreneurship*, 4, 1-63.

Kilcoyne, Patrick. 2001. "High-Tech Occupations by Metropolitan Statistical Area" *Bureau of Labor Statistics.* http://www.bls.gov/oes/2001/tech.pdf

Kirzner, I.M. 1973. *Competition and Entrepreneurship.* Chicago: University of Chicago Press.

Kirzner, I.M. 1979. *Perception, Opportunity, and Profit: Studies in the Theory of Entrepreneurship.* Chicago: University of Chicago Press.

Klein, Peter, G. and Michael L. Cook. 2006. *"T.W. Schultz and the Human-Capital Approach to Entrepreneurship." Review of Agricultural Economics*. 28, 344-350.

Knight, F. 1942. "Profit and Entrepreneurial Functions," *Journal of Economic History.* 2, 126-132.

Koo, Jun. 2005. "How to Analyze the Regional Economy with Occupation Data" *Economic Development Quarterly*, 19, 356-372.

Lazear, Edward P. 2005. "Entrepreneurship," *Journal of Labor Economics*, 23, 649-680.

Lee, Sam Youl, Florida, Richard and Zoltan Aćs. 2004. "Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation." Discussion Papers on Entrepreneurship, Growth and Public Policy. Max Planck Institute for Research into Economic Systems, Group Entrepreneurship, Growth and Public Policy.

Love, James H. 1995. "The Measurement of Entry Rates: Reconsideration and Resolution," *Empirica,* 22, 151-157.

Loveridge, Scott and Denyz Nizalov. 2007. "Operationalizing the Entrepreneurial Pipeline Theory: An Empirical Assessment of the Optimal Size Distribution of Local Firms," *Economic Development Quarterly,* 21, 244-262.

Low, M.B. and I.C. MacMillan. 1988. "Entrepreneurship: Past Research and Future Challenges," *Journal of Management.* 14, 139-161.

Low, Sarah. 2004. "Entrepreneurship Breadth and Depth." *Main Street Economist.* Federal Reserve Bank of Kansas City. September. Available at: http://www.kc.frb.org/RegionalAffairs/mainstreet/MSE_0904.pdf

Low, Sarah, Jason Henderson, and Stephan Weiler. 2005. "Gauging a Region's Entrepreneurship." *Economic Review.* Federal Reserve Bank of Kansas City. Third Quarter.

Low, Sarah A. and Stephan Weiler. 1998. "Risk, Return, and Entrepreneurship," paper presented at the North American meetings of the Regional Science Association International, New York, NY.

Luger, Michael I. and Jun Koo. 2005. "Defining and Tracking Business Start-Ups," *Small Business Economics,* 24, 17-28.

Markusen, Ann Roell. 1985. *Profit Cycles, Oligopoly, and Regional Development. Cambridge, MA:* MIT Press.

Massey, Doreen. 1984. *Spatial Divisions of Labor: Social Structures and the Geography of Production. New York: Methuen.*

McDonald, John and Daniel McMillen. 2006. *Urban Economics and Real Estate: Theory and Policy. New York:* Wiley-Blackwell.

McGranahan, David A. 1999. "Natural Amenities Drive Rural Population Change," Working Paper No. AER-781, United States Department of Agriculture, Economic Research Service.

McGranahan, David A. 2008. "Landscape influence on recent rural migration in the US," *Landscape and Urban Planning,* 85, 228-240.

McGranahan, David and Timothy Wojan. 2007. "Recasting the Creative Class to Examine Growth Processes in Rural and Urban Counties," *Regional Studies,* 41, 197-216.

McGranahan, David, Timothy Wojan, and Dayton Lambert. 2009. "Rural Growth in the Knowledge Economy" *Journal of Economic Geography.* Forthcoming.

McGraw, Thomas K. 2007. *Prophet of Innovation: Joseph Schumpeter and Creative Destruction.* Cambridge, MA: Belknap.

Meager, Niger. 1992. "Does Unemployment lead to self-employment?" *Small Business Economics.* 4, 87-103.

Mueller, Pamela. 2007. "Exploiting Entrepreneurial Opportunities: The Impact of Entrepreneurship on Growth," *The Small Business Economics Journal,* 28, 355-362.

Munn, Johnathan G. 2008. "The Determinants of the Distribution of Innovation: Lessons from Independent Inventors," SRSA Graduate Student Paper prize winner, obtained through personal contact (jmunn@fmarion.edu)

Neumann, George R. and Robert H. Topel. 1991. "Employment Risk, Diversification, and Unemployment," *The Quarterly Journal of Economics,* 106, 1341-1365.

Niosi, Jorge. 2000. "Science-based Industries: a new Schumpeterian Taxonomy," *Technology in Society,* 22, 429-444.

North, Douglass C. 1994. "Economic Performance Through Time," *American Economic Review*, 84(3), 359-68.

Norton, R.D. and J. Rees. 1979. "The Product Cycle and the Spatial Deconcentration of American Manufacturing," *Regional Studies,* 13, 141-151.

Noteboom, B. 1999. "Innovation, Learning and Industrial Organization'" *Cambridge Journal of Economics*, 23, 127-150.

OECD. 2000. *Employment Outlook.* Paris: Organisation for Economic Cooperation and Development

ÓhUallacháin, B. 1999. "'Patent Places: Size Matters'" *Journal of Regional Science*, 39, 613-636.

Orlando, Michael J. and Michael Verba. 2005. "Do Only Big Cities Innovate? Technological Maturity and the Location of Innovation," *Economic Review.* Federal Reserve Bank of Kansas City. Second Quarter.

Parker. Simon C. 1996. "A Time Series Model of Self Employment Under Uncertainty," *Economica*, 63, 469-475.

Parker, Simon C. 2005. "The Economics of Entrepreneurship: What We Know and What We Don't," *Foundations and Trends in Entrepreneurship,* 1, 1-54.

Partridge, Mark D., Dan S. Rickman, Kamar Ali and M. Rose Olfert. 2008. "Lost in Space: Population Dynamics in the American Hinterlands and Small Cities." *Journal of Economic Geography*, 8, 727-757.

Peneder, Michael Rupert. 2008. "Firm Entry and Turnover: the Nexus with Profitability and Growth," *Small Business Economics.* 30, 327-344.

Porter, M.E. and S. Stern. 2003. "The Impact of Location on Global Innovation: Findings from the National Innovative Capacity Index," in P. K. Cornelius (ed.), *The Global Competitiveness Report 2002-2003*, New York: Oxford University Press, pp. 227-253.

Reynolds, Paul, Storey, David and Paul Westhead. 1994. "Cross-national Comparisons of the Variation in New Firm Formation," *Regional Studies,* 28, 443-456.

Reynolds, Paul D. and Richard T. Curtin. 2008. "Business Creation in the United States: Panel Study of Entrepreneurial Dynamics II Initial Assessment", *Foundations and Trends® in Entrepreneurship,* 4, 155-307.

Rocha, Hector and Julian Birkinshaw. 2007. *Foundations and Trends in Entrepreneurship.* Hanover, AM: Now Publishers.

Ross, R. Brent and Randall E. Westgren. 2006. "Economic Returns to Entrepreneurial Behavior," *Journal of Agricultural and Applied Economics,* 38, 403-419.

Saxenian, Annalee. 1994. *Regional Advantage: Culture and Competitive in Silicon Valley and Route 128.* Cambridge, Mass: Harvard University Press.

Say, A. 2001. *A Treatise of Political Economy; or The Production, Distribution and Consumption of Wealth*, Kitchener, Canada: Batoche Books, (First Ed. 1803).

Schiller, Bradley, R. and Philip E. Crewson. 1997. "Entrepreneurial Origins: A Longitudinal Inquiry," *Economic Inquiry,* 35, 523-531.

Schultz, Theodore. 1975. "The Value of the Ability to Deal with Disequilibria," *Journal of Economic Literature,* 3, 827-846

Schumpeter, Joseph A. 1997. *Essays.* London: Transaction Publishers.

Schumpeter, J.A. and R. Opie. 1983. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle, with a new Introduction by John Elliott.* New Brunswick: Transaction Publishers.

Shrestha, Sundar, Goetz, Stephan, J. and Anil Rupasingha. 2007. "Proprietorship Formations and U.S. Job Growth," *The Review of Regional Studies,* 37, 146-168.

Silverman, B.S. 1999. "Technological Resources and the Direction of Corporate Diversification: Toward an Integration of the Resource-Based View and Transaction Cost Economics," *Management Science*, 45, 1109-1124.

Small Business Administration. 2005. *The Innovation-Entrepreneurship Nexus: a National Assessment of Entrepreneurship and Regional Economic Growth and Development.* Office of Advocacy, and Advanced Research Technologies, LLC.

Sorenson, David. 1997. "An Empirical Evaluation of Profit Cycle Theory," *Journal of Regional Science.* Vol. 37, 2. pp. 275-305.

Spelman, William. 2006. "Growth, Stability, and the Urban Portfolio." *Economic Development Quarterly*, 20, 299-315.

Tamasy, Christine. 2006. "Determinants of Regional Entrepreneurship Dynamics in Contemporary Germany: A Conceptual and Empirical Analysis," *Regional Studies,* 40, 365-384.

Trajtenberg, Manuel, Gil Shiff and Ran Melanmed. 2006. "The Name Game: Harnessing Inventors Patents for Research" Working paper 12479. National Bureau for Economic Research.

U.S. Census Bureau. U.S. Census of Population. STF3, 1990. Retrieved August 8, 2008 from http://factfinder.census.gov/home/saff/main.html?_lang=en

Vernon, R. 1966. "International Investment and International Trade in the Product Cycle," *Quarterly Journal of Economics,* 80, 180-207.

von Thünen, J.H. [1826] 1966. *The Isolated State.* London: Pergamon.

Wasylenko, Michael. 1997. "Taxation and economic development: The state of the economic literature," *New England Economic Review,* March/April, 37-52.

Wojan, Timothy R. 2000. Review of Rural Sociology. ERS

Wojan, Timothy R., Dayton M. Lambert, and David A. McGranahan. 2007. "Emoting with their feet: Bohemian attraction to creative milieu," *Journal of Economic Geography,* 1-26.

Wong, P., Ping Ho, Y., and E. Autio. 2005. "Entrepreneurship, Innovation and Economic Growth: Evidence from GEM data," *Small Business Economics*, 24, 335-350.

Yemen, Cory and Michael Lahr. 2008. "Measuring Metropolitan Wage Spillovers of a Redefined Creative Class," paper presented at the North American meetings of the Regional Science Association International, New York, NY.

APPENDIX A: DATA FOR IDENTIFYING ENTREPRENEURIAL INDUSTRIES

I discuss two data sets, establishment births and self employment, which I use to count the number of innovative industry establishments in each county. I argue that these data meet the owner/operator and risk/uncertainty dimensions of entrepreneurship—enabling me to focus on capturing the third dimension, innovation, using the innovative industries. Both datasets are available annually, to best account for cultural and technological change, business cycles, atypical economic events, and maximize flexibility and timeliness of the analysis (Gartner and Shane, 1995).

A.1 ESTABLISHMENT BIRTH DATA

Many researchers use establishment births as a measure of entrepreneurship (Lee et al., 2004; Acs and Mueller, 2008), because establishment births can create growth and increase economic performance (North, 1994). I use single-unit employer establishment births, at the five-digit NAICS industry level, for single-unit establishments in U.S. counties because these data meet the owner/operator and risk/uncertainty bearing attributes of entrepreneurship.

Single-unit employer establishment births meet the owner or operator dimension of entrepreneurship because *a person(s)* must legally establish the firm, be responsible for initial product process or selection, and hiring its first employee. The birth data meet the risk/uncertainty bearing dimension of entrepreneurship because the person who establishes the organization generally has an ownership stake, which in a new establishment, is inherently risky. Risk and uncertainty arise due to the expectation of future sales, profits, establishment success and the risk associated with predicting consumer demand. Single-unit establishment owners bear the up-front costs associated with business operation in the initial phase of operation.

I have access to an establishment dynamics dataset that includes the gross number of establishment births and deaths plus the number of establishments that persisted in each county, in each five-digit industry for multi-unit and single-unit firms. I use single-unit establishments (with a single physical location), that exclude branches, franchises, or subsidiaries of another firm, because single-unit establishments are inherently less entrepreneurial than single-unit establishments due to their independent nature.[14]

Birth data are unsuppressed and, consequently, not publicly available. They are part of a special tabulation of the Statistics of U.S. Businesses series, obtained from the Bureau of the Census, courtesy of the United States Department of Agriculture, Economic Research Service. The publically available version of these data includes the number of establishments in each county for two-digit NAICS; however, this level of aggregation is not suitable for counting innovative industries. Much like Census Bureau economic data, these data exclude establishments with no employees, employees of private households, railroad employees, agricultural production employees, most government employees and professional employer organizations. Where establishments have more than one product, the NAICS codes for their major activity are used.

An establishment birth is defined as an establishment having paid employee(s) in year $t+1$, but not having any paid employees in year t, or not existing in year t. In this dataset, for example, a birth recorded in 2002-2003 indicates the firm had no paid employees in mid-March 2002 and had one or more paid employees in mid-March, 2003. The Census Bureau made careful attempts count only new establishment births by omitting multiple "births" of the same firm that has frequent births/deaths.

Alone, these establishment birth data overestimate entrepreneurship because innovation is ignored. For example, many new establishments replicate existing establishments, e.g., a hair salon or a childcare facility. I overcome this problem by selecting only innovative industry establishments from the employer birth dataset.

A.2 SELF EMPLOYMENT DATA

The self employment rate is a widely used indicator for entrepreneurship because it is readily available, easy to use, and practical (Noteboom, 1999; Schiller and Crewson,

---

[14] Data are based on administrative records; nonsampling errors exist in the data, but precautionary steps were taken by Census Bureau personnel in all phases of collection, processing, and tabulation to minimize the effects of nonsampling errors. Total establishments (births + deaths + persisting establishments) in the 2002-2003 dataset have a 0.9998 correlation with total number of establishments in the 2003 *County Business Patterns* across all U.S. counties. Thus, while the Census Bureau has not disclosed the method for the compilation of the birth dataset, a statistical test shows the datasets not statistically different in number of establishments (Spearman test, Rho=0.993 and P-value <0.000).

1997). I measure self employment with the Census Bureau's Nonemployer Statistics, which include the number of establishments with no paid employees in each county, by six-digit NAICS code.[15]

Regional entrepreneurship researchers often use BEA-REIS nonfarm proprietor data to measure county-level self employment, however, these data are not available at the industry level, so I use the Census' publically available Nonemployer Statistics Series.[16] Nonemployer Statistics include the count of nonemployer establishments, in most industries and at the county-level. Nonemployer establishments are those with no employees who file federal tax Form 1040 (Schedule C), for sole proprietorships, or Form 1065, for partnerships. A diminutive number of incorporated nonemployer establishments are included (Census, 1997) but the Census Bureau tries to eliminate incorporated nonemployer establishments (who use contract employees) by screening out these establishments using an industry-specific gross receipts cutoff. Nonemployer Statistics assigns county of location based upon the tax filing address, the owner's home address, which may outside the county where the business is physically located.

The industries included in Nonemployer Statistics are the same as for the establishment birth data. These data exclude establishments with receipts under $1,000, with the exception of construction industry businesses that are included with receipts over $1. This exclusion omits the smallest of firms, making it a more accurate measure of active small businesses than measures that have no exclusions.

I assume Nonemployer Statistics meet the owner/operator and risk/uncertainty bearing attributes of entrepreneurship, but not the innovation attribute. The self employed are considered owner/operators because there are no employees—thus the owner(s) are responsible for day-to-day operation of the establishment and most nonemployer establishments are proprietorships or partnership that, by definition, meets the ownership attribute. Self employment also includes a degree of risk bearing because uncertainty in business viability and profits is inherent in any private business (Knight, 1942; Cantillon, 1964; Henderson et al., 2006). Self employment overestimates entrepreneurship because

[15] Disclosure problems exist because these data are publicly available, this will be addressed subsequently.
[16] The Census and BEA data are comparable; BEA nonfarm proprietor and Census nonemployer data have a Pearson Correlation of 0.9865 and a difference of means tests rejects the null hypothesis that the two measures are independent, thus, the nonemployer data are not statistically different from the self employment measure that is widely used in the entrepreneurship literature.

many non-innovative firms are included, which Schumpeter argues are no longer entrepreneurial when they have ceased to innovate. Many self employed fail this test because they are lifestyle entrepreneurs who provide replicable services or goods to a local market. By using the self employed in selected, Entrepreneurial Industries, I reduce the overestimation of innovative industries in the self employment rate.

Publically available nonemployer establishment numbers are suppressed for some industries in some counties, unlike the establishment birth data. Most suppressed data are withheld from publication because they would disclose the operations of an individual business, a violation of U.S. Code, Title 13, Section 9. Data are suppressed if a region contains less than three nonemployer establishments in an industry classification. Data are also suppressed for quality purposes. Where industry classification codes are missing, they are imputed and if more than 40 percent of data are for firms with an imputed NAICS code, the data are suppressed because they do not meet publication standards.

I treat the suppressed data as zeros because they represent less than three establishments. The consequence of this, however, is that Entrepreneurial Industries are underestimated—especially in counties with low population or few establishments. Thus, a rural country that has a high level of Computer Service Design and Related Services (NAICS 54151) is *especially* entrepreneurial because the number was high enough to register (three or more establishments). Thus, counties with a non-zero value are especially entrepreneurial when compared to others.

After completion of this dissertation, I would like to propose to the Census Bureau the compilation of the measure using data with no disclosure issues. The raw data will not be made publicly available, rather the nonemployer Entrepreneurial Industries measure.

APPENDIX B: SPATIAL ECONOMETRICS

Controlling for spatial processes in econometric models reduces noise, heterogeneity among units of observation, and improves model fit (Anselin, 1988). Using a given spatial weights matrix, scholars have identified two types of spatial dependence, spatial error (nuisance) dependence and spatial lag (substantive) dependence, and models to control for them.

B.1 SPATIAL WEIGHTS MATRIX

A spatial weights matrix, W, defines spatial unit interaction and is used in spatial econometric models to define neighboring observations. The spatial weights matrix is an $n \times n$ positive matrix that specifies the neighbors for each observation, i.e., it specifies each county's neighboring counties. Each county appears in both row and column and non-zero elements in the matrix indicate a neighbor relation between counties in row *i* and column *j*. By convention, there are no self-neighbors and the weights matrix is row-standardized to facilitate with interpretation and ease computational expense.

There is very little formal guidance for choosing the optimal spatial weights matrix (Anselin, 2008a). The most widely used specification of spatial weights matrices for U.S. counties is the first-order queen contiguity matrix, which specifies a county's neighbors as all counties that are adjacent, in any manner, to the observed county (first-order). Goetz and Rupasingha, however, do not use this W matrix, they use a k=3 nearest-neighbor matrix, which defines neighboring counties as the three closest county centroids to each observation.

Examples of both of these spatial weights matrices are below (Figures B.1, B.2). First-order queen contiguity matrix neighbors for Champaign county (Figure B.1) are A, B, C, D, E, and F—even though F is only contiguous at a vertex; k=3 nearest-neighbors matrix neighbors are A, B, and E because these counties' centroids are closest to the centroid of Champaign county. First-order queen contiguity matrix neighbors for Denver County (Figure B.2) are 1, 2, 4, and 5; k=3 nearest-neighbors matrix neighbors are 3, 4, and 5. Note that the k=3 nearest neighbor matrix neighbors for Denver County include a non-contiguous county because the centroid for county 3 is closer than the centroid of county 1 or county 2.

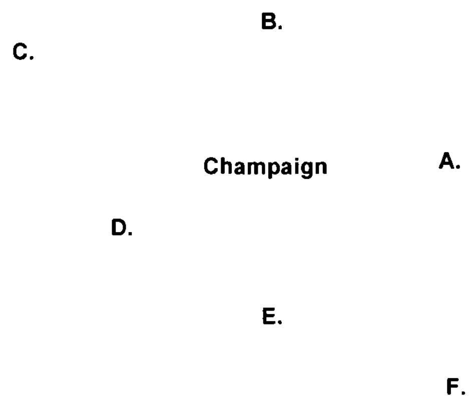

Figure B.1 Champaign County

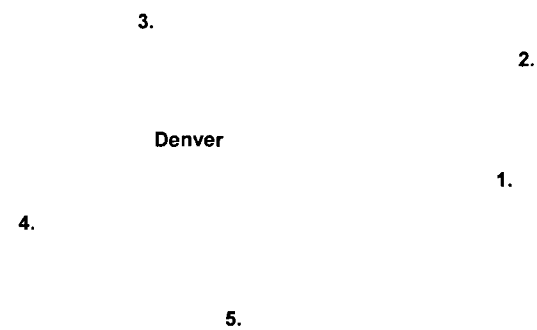

Figure B.2 Denver County

B.2 SPATIAL ERROR PROCESSES

County-level U.S. analysis often suffers from spatial dependence in the error term because county heterogeneity and aerial unit problems create non-spherical disturbances in the error term. *Spatial Error Processes* are spatially correlated disturbances between cross-sectional units and can occur due to omitted spatially correlated variables or the value of adjacent observations moving together due to common or correlated unobservable variables. For this reason, spatial error processes are also known as nuisance errors. I expect county-level models to contain spatial error processes due to heterogeneity of counties, but developing a theory behind implementation of the spatial error model is difficult because the errors are not due to some underlying process—rather a host of micro processes.

Spatial Error Processes can be controlled for in the Spatial Error Model (SEM), which uses the spatial weights matrix to collect non-spherical errors, assuming that remaining errors are spherical, or identically and independently distributed. If spatial error processes are not accounted for, estimates can be inefficient, leading to invalid hypothesis testing. Estimates, however, will not be biased, thus the coefficient sign is not affected.

The SEM model is specified in Equation B.1.

(B.1) $Y = \beta X + \varepsilon$, where, $\varepsilon = \lambda W \varepsilon + u$ and u~i.i.d.

B.3 SPATIAL LAG PROCESSES

Spatial Lag Processes occur due to interaction among neighbors, e.g., copycat behavior, that are due to an underlying spatial process, rather than spatially correlated variables like the spatial error process. For this reason, the spatial lag process creates substantive errors that, when unaccounted for, can lead to biased and inconsistent coefficients, which has the effect of potentially giving the wrong sign on coefficients or leading to invalid hypothesis testing. Thus, spatial lag processes have more dire consequences on estimation than spatial error processes.

Spatial dependence is frequently incorporated into models using the Spatial Autoregressive (SAR) lag model that is not unlike the first order autoregressive model used in time-series analysis. Multiplying the spatial weights matrix by the dependent variable creates a spatially lagged dependent variable with estimated parameter, rho. The SAR model is specified in Equation B.2:

(B.2) $Y = \rho \mathbf{W} Y + \beta X + \varepsilon$, where $\varepsilon \sim$ independent and identically-distributed.

B.4 DIAGNOSITIC TESTS FOR SPATIAL DEPENDENCE

The Lagrange Multiplier (LM) test can indicate whether spatial error and/or spatial lag process are present in OLS regressions. The LM tests for diagnosing spatial dependence have been implemented in various R packages and require constructing a spatial weights matrix, W, and running the OLS regression.

The first step in testing for spatial autocorrelation is to conduct two Lagrange Multiplier (LM) tests, one for a missing spatially lagged dependent variable and a second LM test for error dependence on OLS regressions. If only one of the LM tests fails to reject the null hypothesis, the researcher can stop and proceed with the indicated model. If both LM tests fail to reject the null in favor of the spatial error and spatial lag processes, then Robust LM tests should be conducted and interpreted. The Robust LM test (lag) tests for a missing spatially lagged dependent variable in the possible presence of error dependence, while the Robust LM test (error) tests for error dependence in the possible presence of a missing lagged dependent variable. If both robust tests fail to reject the null hypothesis, the model with the largest coefficient is employed (Florax and Rey, 1995); however, there is academic debate about the appropriateness of this procedure. Goetz and Rupasingha do not follow Florax and Rey's procedure; instead, they use an alternative, higher-order model, which purportedly controls for both spatial error and spatial lag processes.

APPENDIX C: CHAPTER 5 RESULTS

Table C.1 Model 1 and Model 2 Results

	Model 1		Model 2	
Dependent Variable, Y	EI_se_chg		Prop_chg	
	Coeff	Std. Dev.	Coeff	Std. Dev.
Intercept	-0.0018761	0.0009035	0.047598	0.025355
HomeValue	8.5234E-09	9.409E-10	9.282E-09	2.5685E-08
HomeOwn	0.0056	0.0005	-0.0135	0.0136
College	-0.0020	0.0006	-0.0134	0.0159
HS	-0.0004	0.0005	0.0289	0.0148
MedAge	-0.0001	0.0000	0.0007	0.0002
Female	0.0067	0.0015	-0.0400	0.0407
White	0.0003	0.0002	-0.0370	0.0064
WSinc	-0.0001	0.0000	-0.0002	0.0002
WSincGro	0.0001	0.0001	0.0013	0.0020
DeposPop	-0.0047	0.0019	0.0214	0.0502
Unemp	-0.0036	0.0030	-0.0426	0.0805
UnempSq	0.0080	0.0162	-0.0740	0.4387
Nonmetro	-0.0005	0.0001	-0.0031	0.0018
Ag	-0.0003	0.0006	-0.0240	0.0168
Mining	0.0016	0.0012	-0.0425	0.0313
NonDurManu	0.0000	0.0007	0.0595	0.0187
DurManu	0.0003	0.0006	0.0335	0.0156
Trade	0.0000	0.0012	-0.0460	0.0332
Visitor	-0.0002	0.0009	0.0259	0.0251
Services	-0.0015	0.0010	0.0735	0.0278
Amenity	0.0001	0.00002	-0.0014	0.0004
Tax	0.000016	0.000014	-0.0005	0.0004
Level of Y	-0.2692	0.0186	0.0618	0.0143
Lambda	0.06436**		0.10583***	
LogLik	15230.75		5458.198	

Table C.2 Model 3 Results

	Model 3	
Dependent Variable, Y	EI_se_chg	
	Coeff	Std. Dev.
*Intercept^*	-0.001876	0.0009035
HomeValue	8.523E-09	9.409E-10
HomeOwn	0.0056	0.0005
College	-0.0020	0.0006
HS	-0.0004	0.0005
MedAge	-0.0001	0.0000
Female	0.0067	0.0015
White	0.0003	0.0002
DeposPop	-0.0047	0.0019
Nonmetro	-0.0005	0.0001
Ag	-0.0003	0.0006
Mining	0.0016	0.0012
NonDurManu	0.0000	0.0007
DurManu	0.0003	0.0006
Trade	0.0000	0.0012
Visitor	-0.0002	0.0009
Services	-0.0015	0.0010
Amenity	0.0001	0.00002
Level of Y	-0.2692	0.0186
Lambda	0.06436**	
LogLik	15230.75	

^Some explanatory variables were unavailable for 1990s

Table C.3 Model 4, 5, and 6 Results

Dependent Variable	Model 4 El_se/emp Coeff	Std. Dev.	Model 5 EI_birth/pop Coeff	Std. Dev.	Model 6 Prop Coeff	Std. Dev.
Intercept	-0.00537	0.00111	-0.000095	0.000032	0.55504	0.04133
HomeValue	1.49E-08	1.14E-09	3.28E-10	3.33E-11	3.93E-07	4.26E-08
HomeOwn	0.0075	0.0006	0.000013	0.000018	0.1730	0.0226
College	0.0015	0.0007	0.000326	0.000021	0.0976	0.0269
HS	-0.0003	0.0007	-0.000018	0.000020	0.0188	0.0250
MedAge	-0.0001	0.0000	0.0000010	0.0000003	0.0050	0.0004
Female	0.0125	0.0018	0.000032	0.000052	-0.8228	0.0665
White	0.0015	0.0003	0.000006	0.000008	-0.0469	0.0109
WSinc	0.0000	0.0000	-0.0000002	0.0000002	-0.0106	0.0003
WSincGro	0.0004	0.0001	0.000017	0.000003	-0.0128	0.0033
DeposPop	-0.0073	0.0023	0.000203	0.000067	0.0395	0.0852
Unemp	-0.0035	0.0037	-0.000126	0.000106	-0.2661	0.1361
UnempSq	0.0032	0.0199	-0.000035	0.000582	0.9151	0.7431
Nonmetro	-0.0012	0.0001	0.000002	0.000002	-0.0203	0.0030
Ag	-0.0023	0.0008	0.000088	0.000022	0.0243	0.0285
Mining	0.0016	0.0014	0.000040	0.000041	-0.0067	0.0530
NonDurManu	0.0004	0.0008	0.000015	0.000025	-0.0858	0.0315
DurManu	0.0000	0.0007	-0.000005	0.000021	0.0079	0.0265
Trade	0.0014	0.0015	0.000010	0.000044	-0.1404	0.0562
Visitor	0.0017	0.0011	0.000118	0.000033	0.0892	0.0425
Services	-0.0052	0.0013	0.000039	0.000037	-0.0932	0.0468
Amenity	0.0001	0.00002	0.000001	0.000001	-0.0007	0.0007
Tax	0.00003	0.00002	0.000001	0.000001	0.0010	0.0007
Lambda	0.1698***		0		0.07275**	
LogLik	14615.14		25081.61		3895.73	

Table C.4 Model 7 and 8 Results

	Model 7		Model 8	
Dependent Variable, Y	ST2		STP	
	Coeff	Std. Dev.	Coeff	Std. Dev.
Intercept	-0.0185	0.0021	-0.0043	0.0035
HomeValue	3.98E-08	2.18E-09	4.78E-08	3.58E-09
HomeOwn	0.0140	0.0012	0.0257	0.0019
College	0.0227	0.0014	-0.0117	0.0023
HS	0.0011	0.0013	-0.0013	0.0021
MedAge	-0.0001	0.0000	-0.0002	0.0000
Female	0.0272	0.0034	0.0188	0.0056
White	0.0028	0.0006	0.0070	0.0009
WSinc	-0.0002	0.0000	-0.0005	0.0000
WSincGro	0.0006	0.0002	0.0014	0.0003
DeposPop	-0.0189	0.0043	-0.0033	0.0071
Unemp	0.0061	0.0070	0.0017	0.0115
UnempSq	-0.0082	0.0380	-0.0041	0.0622
Nonmetro	-0.0020	0.0002	-0.0020	0.0003
Ag	-0.0037	0.0015	-0.0114	0.0024
Mining	0.0017	0.0027	-0.0027	0.0044
NonDurManu	-0.0006	0.0016	0.0056	0.0027
DurManu	-0.0002	0.0014	0.0019	0.0022
Trade	0.0016	0.0029	0.0015	0.0047
Visitor	0.0045	0.0022	-0.0024	0.0036
Services	-0.0078	0.0024	-0.0179	0.0039
Amenity	0.0004	0.00004	0.0005	0.0001
Tax	0.000093	0.000036	-0.000043	0.000060
Lambda	0.14706***		0.19918***	
LogLik	12701.2		11237.57	

APPENDIX D: CHAPTER 6 RESULTS

Table D.1 Full Results, Table 6.3, ESHIP=*EI_SE/EMP*

	EI_se/emp _Y=JobChg		EI_se/emp _Y=PopChg		EI_se/emp _Y=EstabChg	
	Coeff	Std. Dev.	Coeff	Std. Dev.	Coeff	Std. Dev.
Intercept	-0.5730	0.0697	-0.6264	0.0450	-0.5595	0.0640
Creative	0.0075	0.0031	0.0095	0.0020	0.0222	0.0027
E'ship	0.0134	0.0022	0.0174	0.0014	0.0046	0.0019
Creative X e'ship	-0.0013	0.0013	-0.0043	0.0008	-0.0057	0.0009
HS	-0.1160	0.0379	-0.1109	0.0244	-0.0689	0.0346
College	-0.0416	0.0306	-0.0297	0.0198	-0.0187	0.0281
EmpRate	0.0379	0.0267	0.0104	0.0173	0.0000	0.0247
MedInc	5.85E-07	3.28E-07	2.71E-07	2.12E-07	3.90E-07	3.00E-07
PopDen	-2.35E-06	1.05E-06	-1.85E-06	6.76E-07	4.59E-07	9.67E-07
Commute	0.1108	0.0110	0.0519	0.0071	0.0380	0.0100
Metro	0.0075	0.0043	0.0111	0.0028	0.0075	0.0039
Ag	0.0401	0.0365	-0.0054	0.0236	0.0417	0.0334
Mining	-0.0686	0.0669	-0.1176	0.0432	-0.1080	0.0612
NonDurManu	-0.1145	0.0414	-0.0300	0.0268	-0.0001	0.0380
DurManu	-0.1650	0.0346	-0.0637	0.0224	-0.1299	0.0317
Trade	0.0647	0.0709	0.1119	0.0456	0.1705	0.0644
Visitor	-0.0016	0.0717	0.0343	0.0463	0.0632	0.0653
Services	0.0207	0.0602	-0.0030	0.0388	0.0717	0.0548
Pop8-17	0.3625	0.1044	-0.1076	0.0677	0.3390	0.0969
Pop62	0.1711	0.0506	-0.2696	0.0327	-0.0201	0.0466
PctBlack	-0.0507	0.0144	-0.0368	0.0094	-0.0409	0.0131
PctNA	0.0063	0.0281	0.0233	0.0182	-0.1110	0.0258
PctHis	0.0466	0.0156	0.0178	0.0102	-0.0309	0.0145
Military	0.1817	0.0667	-0.0727	0.0430	0.2676	0.0609
CollegePop	0.0530	0.0336	0.0081	0.0216	-0.0182	0.0309
OutAmen	0.0162	0.0026	0.0105	0.0017	0.0136	0.0024
PubLand	-0.0023	0.0023	-0.0038	0.0015	0.0003	0.0021
EstabChg90	0.4378	0.0569	0.1098	0.0367	0.2579	0.0522
JobChg90	0.1090	0.0125	0.1472	0.0080	0.1053	0.0114
PopChg90	-0.00014	0.0001	0.00002	0.0001	0.00005	0.0001
Lambda	0.1436***		0.0871***		0.2006***	
LogLik	4545.79		3249.72		3507.62	

Table D.2 Full Results, Table 6.3, ESHIP=*EI_BIRTH/POP*

	EI_birth/pop Y=JobChg Coeff	Std. Dev.	EI_birth/pop Y=PopChg Coeff	Std. Dev.	EI_birth/pop Y=EstabChg Coeff	Std. Dev.
Intercept	-0.5578	0.0699	-0.6142	0.0458	-0.5595	0.0640
Creative	0.0195	0.0029	0.0188	0.0019	0.0222	0.0027
E'ship	0.0020	0.0021	0.0043	0.0014	0.0046	0.0019
Creative X e'ship	-0.0061	0.0010	-0.0058	0.0007	-0.0057	0.0009
HS	-0.0982	0.0380	-0.0987	0.0248	-0.0689	0.0346
College	-0.0675	0.0306	-0.0546	0.0201	-0.0187	0.0281
EmpRate	0.0298	0.0268	0.0005	0.0176	0.0000	0.0247
MedInc	6.56E-07	3.28E-07	3.32E-07	2.15E-07	3.90E-07	3.00E-07
PopDen	-1.06E-06	1.06E-06	-5.09E-07	6.95E-07	4.59E-07	9.67E-07
Commute	0.1184	0.0109	0.0562	0.0072	0.0380	0.0100
Metro	0.0064	0.0043	0.0127	0.0028	0.0075	0.0039
Ag	0.0417	0.0365	-0.0133	0.0239	0.0417	0.0334
Mining	-0.0653	0.0669	-0.1200	0.0438	-0.1080	0.0612
NonDurManu	-0.1036	0.0415	-0.0157	0.0272	-0.0001	0.0380
DurManu	-0.1577	0.0346	-0.0546	0.0227	-0.1299	0.0317
Trade	0.0689	0.0708	0.1263	0.0463	0.1705	0.0644
Visitor	0.0369	0.0716	0.0611	0.0468	0.0632	0.0653
Services	0.0083	0.0601	-0.0115	0.0393	0.0717	0.0548
Pop8-17	0.2936	0.1051	-0.1767	0.0690	0.3390	0.0969
Pop62	0.1122	0.0508	-0.3297	0.0333	-0.0201	0.0466
PctBlack	-0.0629	0.0142	-0.0541	0.0093	-0.0409	0.0131
PctNA	0.0065	0.0281	0.0220	0.0184	-0.1110	0.0258
PctHis	0.0480	0.0156	0.0167	0.0103	-0.0309	0.0145
Military	0.1457	0.0669	-0.0983	0.0437	0.2676	0.0609
CollegePop	-0.0008	0.0339	-0.0310	0.0222	-0.0182	0.0309
OutAmen	0.0179	0.0026	0.0132	0.0017	0.0136	0.0024
PubLand	-0.0038	0.0022	-0.0059	0.0015	0.0003	0.0021
EstabChg90	0.4369	0.0572	0.1020	0.0374	0.2579	0.0522
JobChg90	0.1149	0.0124	0.1536	0.0081	0.1053	0.0114
PopChg90	-0.0001	0.0001	0.0001	0.0001	0.00005	0.0001
Lambda	0.0931***		0.1311***		0.1984***	
LogLik	3248.44		4504.59		3511.62	

Table D.3 Full Results, Table 6.4, ESHIP=*SELFEMP*

	SelfEmp Y=JobsChg		SelfEmp Y=PopChg		SelfEmp Y=EstabChg	
	Coeff	Std. Dev.	Coeff	Std. Dev.	Coeff	Std. Dev.
Intercept	-0.4611	0.0704	-0.6926	0.0468	-0.6068	0.0654
Creative	0.0153	0.0026	0.0133	0.0017	0.0173	0.0024
E'ship	0.0194	0.0021	-0.0074	0.0014	-0.0029	0.0019
Creative X e'ship	0.0018	0.0017	0.0014	0.0011	0.0001	0.0015
HS	-0.1395	0.0377	-0.1015	0.0250	-0.0775	0.0348
College	-0.0500	0.0303	-0.0457	0.0201	-0.0075	0.0282
EmpRate	0.0259	0.0265	0.0090	0.0176	0.0073	0.0247
MedInc	7.58E-07	3.26E-07	2.37E-07	2.16E-07	3.26E-07	3.03E-07
PopDen	-1.92E-06	1.05E-06	-1.26E-06	6.99E-07	-3.90E-07	9.73E-07
Commute	0.0644	0.0119	0.0668	0.0079	0.0375	0.0110
Metro	0.0141	0.0042	0.0165	0.0028	0.0115	0.0039
Ag	0.0026	0.0362	-0.0172	0.0241	0.0330	0.0336
Mining	-0.0912	0.0664	-0.1326	0.0441	-0.1210	0.0615
NonDurManu	-0.0810	0.0413	-0.0374	0.0274	-0.0134	0.0383
DurManu	-0.1509	0.0344	-0.0686	0.0228	-0.1409	0.0319
Trade	0.0949	0.0701	0.1317	0.0466	0.1796	0.0648
Visitor	0.0108	0.0709	0.0371	0.0471	0.0403	0.0656
Services	0.0236	0.0596	-0.0289	0.0396	0.0600	0.0552
Pop8-17	0.2098	0.1048	-0.0859	0.0697	0.3974	0.0978
Pop62	-0.0227	0.0531	-0.2352	0.0353	0.0299	0.0493
PctBlack	-0.0359	0.0145	-0.0700	0.0097	-0.0486	0.0136
PctNA	0.0175	0.0280	0.0081	0.0186	-0.1196	0.0260
PctHis	0.0418	0.0155	0.0122	0.0103	-0.0357	0.0146
Military	0.1691	0.0660	-0.0765	0.0439	0.2869	0.0611
CollegePop	0.0401	0.0330	0.0065	0.0219	0.0137	0.0305
OutAmen	0.0162	0.0026	0.0148	0.0017	0.0150	0.0024
PubLand	-0.0059	0.0022	-0.0059	0.0015	-0.0001	0.0021
EstabChg90	0.3607	0.0585	0.1198	0.0389	0.2733	0.0541
JobChg90	0.1025	0.0124	0.1606	0.0082	0.1087	0.0115
PopChg90	-0.00012	0.0001	0.00004	0.0001	0.00002	0.0001
Lambda	0.1157***		0.1216***		0.1902***	
LogLik	3273.72		4484.69		3494.27	

Table D.4 Full Results, Table 6.4, ESHIP=*ESTAB/EMP*

	Estab/Emp Y=JobChg		Estab/Emp Y=PopChg		Estab/Emp Y=EstabChg	
	Coeff	Std. Dev.	Coeff	Std. Dev.	Coeff	Std. Dev.
Intercept	-0.6571	0.0703	-0.6370	0.0467	-0.5878	0.0646
Creative	0.0108	0.0026	0.0140	0.0018	0.0158	0.0024
E'ship	0.0153	0.0021	-0.0011	0.0014	-0.0010	0.0019
Creative X e'ship	0.0003	0.0016	-0.0004	0.0011	-0.0076	0.0015
HS	-0.1273	0.0379	-0.1103	0.0252	-0.0698	0.0347
College	-0.0594	0.0304	-0.0424	0.0203	-0.0094	0.0281
EmpRate	0.0179	0.0267	0.0073	0.0178	0.0097	0.0247
MedInc	7.03E-07	3.27E-07	2.88E-07	2.18E-07	3.46E-07	3.01E-07
PopDen	-1.95E-06	1.05E-06	-1.44E-06	6.95E-07	-7.43E-07	9.58E-07
Commute	0.1062	0.0108	0.0493	0.0072	0.0339	0.0099
Metro	0.0145	0.0042	0.0168	0.0028	0.0102	0.0039
Ag	0.0068	0.0364	-0.0245	0.0242	0.0380	0.0335
Mining	-0.0600	0.0668	-0.1360	0.0444	-0.1154	0.0613
NonDurManu	-0.1006	0.0414	-0.0249	0.0275	-0.0048	0.0380
DurManu	-0.1554	0.0346	-0.0634	0.0230	-0.1312	0.0318
Trade	0.0654	0.0706	0.1392	0.0469	0.1876	0.0646
Visitor	0.0145	0.0713	0.0365	0.0473	0.0423	0.0653
Services	-0.0096	0.0599	-0.0195	0.0398	0.0703	0.0549
Pop8-17	0.2675	0.1051	-0.1319	0.0700	0.3369	0.0972
Pop62	0.0082	0.0541	-0.2948	0.0360	-0.0220	0.0498
PctBlack	-0.0591	0.0142	-0.0575	0.0095	-0.0447	0.0132
PctNA	0.0140	0.0282	0.0162	0.0187	-0.1072	0.0259
PctHis	0.0453	0.0156	0.0120	0.0104	-0.0330	0.0145
Military	0.2245	0.0669	-0.0823	0.0444	0.2766	0.0613
CollegePop	0.0675	0.0335	-0.0017	0.0223	-0.0106	0.0307
OutAmen	0.0182	0.0026	0.0143	0.0017	0.0164	0.0024
PubLand	-0.0055	0.0022	-0.0064	0.0015	-0.0001	0.0021
EstabChg90	0.2473	0.0632	0.1238	0.0420	0.3164	0.0579
JobChg90	0.1417	0.0129	0.1529	0.0086	0.1081	0.0119
PopChg90	-0.0001	0.0001	0.0000	0.0001	-0.00002	0.0001
Lambda	0.0983***		0.1805***		0.1918***	
LogLik	3256.53		3700.60		3506.77	

Table D.5 Full Results, Table 6.5, ESHIP=*ST2*

	ST2 Y=JobChg Coeff	Std. Dev.	ST2 Y=PopChg Coeff	Std. Dev.	ST2 Y=EstabChg Coeff	Std. Dev.
Intercept	-0.5480	0.0697	-0.5932	0.0449	-0.5491	0.0637
Creative	0.0060	0.0032	0.0104	0.0021	0.0153	0.0030
E'ship	0.0194	0.0026	0.0209	0.0017	0.0159	0.0024
Creative X e'ship	-0.0040	0.0012	-0.0075	0.0008	-0.0060	0.0011
HS	-0.1093	0.0378	-0.1000	0.0243	-0.0721	0.0344
College	-0.0559	0.0305	-0.0491	0.0197	-0.0129	0.0280
EmpRate	0.0219	0.0266	-0.0084	0.0172	-0.0051	0.0245
MedInc	5.44E-07	3.27E-07	2.32E-07	2.11E-07	3.14E-07	3.00E-07
PopDen	-2.55E-06	1.05E-06	-1.68E-06	6.74E-07	-6.10E-07	9.54E-07
Commute	0.1103	0.0110	0.0529	0.0071	0.0343	0.0100
Metro	0.0063	0.0043	0.0097	0.0028	0.0060	0.0039
Ag	0.0378	0.0364	-0.0066	0.0234	0.0442	0.0333
Mining	-0.0610	0.0667	-0.1082	0.0429	-0.1032	0.0609
NonDurManu	-0.1120	0.0413	-0.0260	0.0266	-0.0095	0.0378
DurManu	-0.1637	0.0345	-0.0599	0.0222	-0.1361	0.0316
Trade	0.0630	0.0706	0.1093	0.0453	0.1585	0.0642
Visitor	0.0118	0.0715	0.0547	0.0459	0.0546	0.0651
Services	0.0117	0.0599	-0.0143	0.0385	0.0664	0.0546
Pop8-17	0.3650	0.1040	-0.1150	0.0673	0.3926	0.0959
Pop62	0.1539	0.0502	-0.2965	0.0324	0.0118	0.0460
PctBlack	-0.0502	0.0143	-0.0401	0.0093	-0.0305	0.0132
PctNA	0.0020	0.0280	0.0183	0.0181	-0.1143	0.0256
PctHis	0.0425	0.0156	0.0133	0.0101	-0.0340	0.0145
Military	0.1761	0.0665	-0.0853	0.0427	0.2812	0.0606
CollegePop	0.0247	0.0332	-0.0247	0.0213	-0.0080	0.0302
OutAmen	0.0135	0.0027	0.0084	0.0017	0.0102	0.0025
PubLand	-0.0028	0.0023	-0.0047	0.0015	0.0010	0.0021
EstabChg90	0.4258	0.0567	0.0921	0.0364	0.2531	0.0517
JobChg90	0.1100	0.0124	0.1484	0.0080	0.1016	0.0113
PopChg90	-0.0001	0.0001	0.0000	0.0001	0.00003	0.0001
Lambda	0.0997***		0.16225***		0.2106***	
LogLik	3259.43		4563.76		3522.77	

Table D.6 Results, Table 6.5, ESHIP=*STP*

	STP Y=JobChg Coeff	Std. Dev.	STP Y=PopChg Coeff	Std. Dev.	STP Y=EstabChg Coeff	Std. Dev.
Intercept	-0.5755	0.0697	-0.6279	0.0450	-0.5782	0.0639
Creative	0.0079	0.0031	0.0097	0.0020	0.0165	0.0028
E'ship	0.0139	0.0023	0.0182	0.0015	0.0105	0.0021
Creative X e'ship	-0.0020	0.0012	-0.0049	0.0008	-0.0039	0.0011
HS	-0.1148	0.0379	-0.1092	0.0244	-0.0789	0.0346
College	-0.0432	0.0306	-0.0313	0.0198	-0.0024	0.0281
EmpRate	0.0365	0.0267	0.0091	0.0173	0.0071	0.0246
MedInc	5.82E-07	3.28E-07	2.64E-07	2.12E-07	3.45E-07	3.01E-07
PopDen	-2.47E-06	1.05E-06	-1.93E-06	6.75E-07	-7.21E-07	9.56E-07
Commute	0.1124	0.0110	0.0533	0.0071	0.0352	0.0101
Metro	0.0070	0.0043	0.0104	0.0028	0.0075	0.0039
Ag	0.0400	0.0365	-0.0052	0.0236	0.0425	0.0335
Mining	-0.0704	0.0669	-0.1193	0.0432	-0.1132	0.0612
NonDurManu	-0.1145	0.0414	-0.0293	0.0268	-0.0118	0.0380
DurManu	-0.1661	0.0346	-0.0645	0.0223	-0.1400	0.0317
Trade	0.0625	0.0709	0.1091	0.0456	0.1625	0.0645
Visitor	0.0007	0.0718	0.0365	0.0462	0.0457	0.0655
Services	0.0175	0.0602	-0.0054	0.0388	0.0676	0.0550
Pop8-17	0.3650	0.1044	-0.1075	0.0677	0.3969	0.0964
Pop62	0.1720	0.0506	-0.2695	0.0327	0.0250	0.0464
PctBlack	-0.0508	0.0144	-0.0370	0.0094	-0.0318	0.0133
PctNA	0.0067	0.0281	0.0240	0.0182	-0.1117	0.0258
PctHis	0.0466	0.0156	0.0179	0.0102	-0.0315	0.0145
Military	0.1812	0.0667	-0.0731	0.0430	0.2875	0.0609
CollegePop	0.0506	0.0336	0.0057	0.0216	0.0100	0.0306
OutAmen	0.0161	0.0026	0.0103	0.0017	0.0123	0.0024
PubLand	-0.0024	0.0023	-0.0039	0.0015	0.0011	0.0021
EstabChg90	0.4413	0.0569	0.1129	0.0367	0.2675	0.0519
JobChg90	0.1100	0.0125	0.1481	0.0080	0.1029	0.0114
PopChg90	-0.0001	0.0001	0.0000	0.0001	0.00002	0.0001
Lambda	0.0871***		0.1463***		0.2017***	
LogLik	3249.40		4547.76		3508.38	

CURRICULUM VITAE

Sarah A. Low

707 N. Wayne St. #302, Arlington, VA 22201 — 765.430.4454 (cell)
sarahalow@gmail.com — 202.694.5603 (office)

Education

University of Illinois at Urbana-Champaign,
Dept. of Agricultural & Consumer Economics — Urbana, IL — PhD. Oct., 2009
Primary Field: Regional Economics & Public Policy
Secondary Field: Environmental & Natural Resource Economics
Dissertation: *Defining and Measuring Entrepreneurship for Regional Research: A New Approach*

Purdue University,
Dept. of Agricultural Economics — West Lafayette, IN — M.S. May, 2004
Major: Agricultural Economics, focus in regional economics
Thesis: *Employment Growth and Fiscal Impact Analysis of a Job Creation Tax Credit for the Indiana Enterprise Zones*

Iowa State University, — Ames, IA — B.S. May, 2002 with Honors & Distinction
Dept. of Sociology
Major: Public Service and Administration in Agriculture
Minor: Entrepreneurial Studies
Honors Thesis: *Text Preparation and Teaching of Sociology 130: Rural Organizations and institutions*

Research Experience

***Graduate Research Assistant*, Dr. Andrew M. Isserman**
Regional Economics and Public Policy (REAP)
Dept. of Agricultural & Consumer Economics, University of Illinois at Urbana-Champaign
Urbana, IL — 8/06-7/09
-Entrepreneurship as an economic development tool
-Spatial analysis of U.S. ethanol industry
-Economic impact analysis (input-output) of ethanol plants

Economist (Student Temporary Employment Program)
U.S. Dept. of Agriculture, Economic Research Service
Washington, DC — 8/08-5/09
-Defining and measuring entrepreneurship
-Rural broadband availability, infrastructure & impacts

Research Associate II
Federal Reserve Bank of Kansas City, Center for the Study of Rural America
Kansas City, MO — 2/04-8/06
-Wrote articles for monthly policy brief, *Main Street Economist*
-Developed county-level Regional Asset Indicator Series, data and articles
-Research on the spatial measures of high-value entrepreneurship and innovation
-Work on county-level growth modeling using asset indicators as explanatory variables
-Assist with public information duties, including giving speeches
-Responsible for hiring and supervision of graduate student interns

***Graduate Research Assistant*, Dr. Kevin T. McNamara**
Purdue University, Department of Agricultural Economics
West Lafayette, IN 8/02-1/04
-Determined fiscal impact of proposed tax incentives for enterprise zones
-Conducted research on Indiana's food manufacturing sector and farm size and structure

Fiscal Analyst Intern, Office of Fiscal & Management Analysis
Indiana Legislative Services Agency
Indianapolis, IN 12/02-12/03
-Research on state enterprise zone programs
Designed and implemented business survey, Geocoded ES202 data to develop business database, Trained seven people in geographic information systems (GIS), Presented results to stakeholders
-Conducted research on fiscal impact of Streamlined Sales Tax Project (SSTP) adoption

Intern to Marketing and Food Research Program Leader
Leopold Center for Sustainable Agriculture
Ames, IA 5/01-4/02
-Researched niche and direct marketing of agricultural products as tools to increase farm income
-Contributed to grant report writing and local food awareness
-Conducted focus group discussion panel at food use seminar

Teaching Experience

Regional Economics and Public Policy, University of Illinois at Urbana-Champaign
Urbana, IL May 19-21, 2008
Co-Instructor: Building Skills for Outreach Research, University of Illinois Short Course
-Developed ArcGIS curriculum for short course on conducting applied research
-Tutored 14 extension specialists in database management, GIS, and conducting regional research

Dept. of Agricultural and Consumer Economics (ACE), University of Illinois
Urbana, IL Spring, 2008
Instructor: ACE 499-RD, Computer Applications and Data Analysis for Rural Development
-Developed curriculum for undergraduate course in rural development, data analysis, and GIS
-Taught 25 junior and senior undergraduates from Departments of ACE and ECON, Business and Liberal Arts colleges
-Instructor of record for 2 credit hour course
-Overall Instructor Rank (student evaluations): 4.3/5.0

Department of Sociology, Dr. Steve Padgitt, Iowa State University
Ames, IA 8/00-5/02
Teaching assistant: SOC 130, Rural Institutions & Organizations
-Developed curriculum and Led two weekly recitation (discussion) sections

Department of Sociology, Dr. Paul Lasley, Iowa State University
Ames, IA Spring 2001, 2002
Teaching assistant: SOC 325, Agriculture in Transition
-Write, administer and grade exams, take lecture notes and attendance

Boilermaker TaeKwonDo, Purdue University
West Lafayette, IN 1/2003-5/2004
Assistant Instructor: Beginner and Color Belt TaeKwonDo
-Self Defense and Sparring
-Rank: 1st Degree Black Belt, Kukkiwon

Publications

Professional Publications

Ahearn, M.C., M. Kilkenney, & S.A. Low. *Forthcoming*. "Trends and Impacts of Rural School Finance." *Review of Agricultural Economics.*

Low, S.A. & A.M. Isserman. 2009. 'Ethanol and the Local Economy: Industry Trends, Location Factors, Economic Impacts, and Risks." *Economic Development Quarterly*. February.

Henderson, J., S.A. Low, & S. Weiler. 2007. Book Chapter, Ch. 5. "The Drivers of Regional Entrepreneurship in Rural and Metro Areas." *Entrepreneurship and Local Economic Development* edited by Norman Walzer. Lexington Books.

Low, S.A. 2006. Book Review, "Entrepreneurship in the Region," edited by Michael Fritsch and Jürgen Schmude. *Papers in Regional Science*. June, p. 331-333

Low, S.A., J. Henderson & S. Weiler. 2005. Gauging a Region's Entrepreneurial Potential. *Economic Review*, Federal Reserve Bank of Kansas City, pp.61-89, Third Quarter.

Policy Publications

Low, S. & P. Stenberg. 2009. "Rural Broadband At A Glance." United States Department of Agriculture, Economic Research Service. March.

Low, S. & A. Isserman. 2007. Chapter 5. "Ethanol and the Local Economy." Corn-Based Ethanol in Illinois and the U.S.: A Report from the Department of Agricultural and Consumer Economics, University of Illinois. November.

Henderson, J. & S. Low. 2006. "Obesity: America's Economic Epidemic," *Main Street Economist*, Federal Reserve Bank of Kansas City, Issue #2.

Low, S. 2005. "Regional Asset Indicators: The Wealth of Regions," *Main Street Economist*, Federal Reserve Bank of Kansas City, September.

Low, S. 2005. "Regional Asset Indicators: Bank Deposit Depth and Evolution," *Main Street Economist,* Federal Reserve Bank of Kansas City, April.

Low, S. 2005. "New Federalism—Problems or Opportunities for Rural America?" *New Issues on the Rural Horizon*, Center for the Study of Rural America, Annual Report.

Low, S. 2004 "An Economy of Regions." *Building Rural Prosperity in Regions: The Road Less Traveled*, Center for the Study of Rural America, Annual Report.

Low, S. 2004. "Regional Asset Indicators: Entrepreneurship Breadth and Depth," *Main Street Economist*, Federal Reserve Bank of Kansas City, September.

Low, S. & K. McNamara. 2003. "Indiana Food Industry: Size, Income, and Employment Effects," Purdue University, Staff Paper #2003-03, February.

Low, S. & K. McNamara. 2003. "Indiana Production Agriculture: Size, Structure, and Income Effects," Purdue University, Staff Paper #2003-04, February.

Academic Papers Presented

Measuring and Defining Entrepreneurship. North American meetings of the Regional Science Association International. Brooklyn, NY. November 2008.

Risk, Returns, and Entrepreneurship. North American meetings of the Regional Science Association International. Brooklyn, NY. November 2008.

Ethanol: Implications for Communities. Applied and Agricultural Economics Association. Orlando, FL. July 2008.

Measuring and Defining Entrepreneurship in Rural Regions. PREPARE/European Regional Science Association Summer School. University of Pécs, Pécs, Hungary. July 2008.

Regional Employment, Risk, Returns, and Entrepreneurship. Mid-Continent Regional Science Association. Colorado Springs, CO. June 2008.

Obesity in Rural America: Reasons, Impacts, and Solutions for Communities. Southern Regional Science Association. Arlington, VA. March 2008.

Ethanol: Implications for Rural Communities. North American meetings of the Regional Science Association International. Savannah, GA. November 2007.

Innovation and Entrepreneurship: Framing the Question in Rural Space. Frameworks for Entrepreneurship Research in Food, Agriculture and Rural Development Workshop. Kauffman Foundation, Kansas City, MO. October 2007.

Ethanol, Economic Growth, and Public Policy: Implications for Rural Regions, Southern Regional Science Association, Charleston, SC. March 2007.

Innovation and Entrepreneurship in Rural America: Empirical Effects on Economic Growth, North American meetings of the Regional Science Association International. Toronto, OT. November 2006.

Measures of Entrepreneurship Breadth and Depth over Space and Time, Western Regional Science Association, Santa Fe, NM, February 2005

Gauging A Region's Entrepreneurial Potential, University of Missouri, Contracting and Organization Research Institute Seminar, December 2005

Economic Growth, Asset Indicators, and Entrepreneurship, North American meetings of the Regional Science Association International., Las Vegas, NV, November 2005

Gauging A Region's Entrepreneurial Potential, American Agricultural Economics Association, Providence, RI, July 2005

Measures of Entrepreneurship Breadth, Income, and Value Added Over Space and Time, Western Economics Association International, San Francisco, CA, July 2005

High Value Entrepreneurship, Banking Deposit Depth and Regional Growth, Southern Regional Science Association, Washington, DC, April 2005

Where are the Entrepreneurs? Examining the Location of Proprietor Breadth and Depth, North American meetings of the Regional Science Association International. Seattle, WA, November 2004

The Indiana Enterprise Zone Program: Fiscal Impact of a Job Creation Tax Credit, American Agricultural Economics Association, Denver, CO, July 2004

State Fiscal Impact of Indiana Enterprise Zone Job Creation Tax Credit, Southern Regional Science Association, New Orleans, LA, March 2004

Utilizing GIS to Analyze Economic Regions, Indiana Geographic Information System Conference, Indianapolis, IN, March 2004

Professional Presentations

Ethanol: Implications for Illinois. United Counties Council of Illinois, Annual Meeting. Galena, IL. July 2007.

Discussant: Immigrants, Farm-Related Industries, and Communities, *Immigration Reform: Implications for Farmers, Farm-Workers and US Agriculture,* University of California, Washington DC, June 2006

Midwest Demographics and Agriculture, Rotary District 9970 Annual Convention, Hokitika, New Zealand, April 2006

Local Entrepreneurship: One Size Does Not Fit All, Economic Development Administration: University Center Conference, Reno, NV, August 2005

New Opportunities for Washington's Agricultural Economy, Washington State Bankers Association, May 2005
New Economic Opportunities for Northwest Kansas, Northwest Kansas Planning and Development Commission, December 2004
New Opportunities for Main Street Nebraska, North Platte Rotary Club, June 2004

Honors

Selected Participant. National Bureau of Economic Research, Entrepreneurship Research Boot Camp. Cambridge, MA. July 2009
Selected Young Researcher Participant. European Regional Science Association, Spatial Econometrics and Spatial Computable General Equilibrium Modeling Workshop. Pécs, Hungary. July 2008.
Springer-Verlag Prize, best paper by an early-career-stage scholar, Western Regional Science Association, February 2006
Group Study Exchange Team Member, work with farmers, economists, and Rotarians in New Zealand to learn about New Zealand agricultural and economic development policy and practices, March-April 2006
"Red" Barron All-University Senior Leadership Award, presented for outstanding levels of undergraduate leadership, Iowa State University, April 2002
University Honors Program-Full Member, Iowa State
Dean's List-Iowa State University 2000-2002

Professional Activities

Member: American Agricultural Economics Association
Graduate Student Section Officer, Membership. 2007-2008
-assist with execution of GSS Track Sessions and competitions
-responsible for member communication
-publicity of section activities at annual meeting
North American Regional Science Council
Southern Regional Science Association
Manuscript Referee: Economic Development Quarterly, Review of Regional Studies, Environment and Planning
Peer Reviewer: American Agricultural Economics Association

Other Leadership Experience

Ambassador Coordinator, United Way Steering Committee, Federal Reserve Bank, 2006
Volunteer of the Month, Community Investment Program, June 2006
Chief of Staff, Government of the Student Body, Iowa State University, 2001-2002
-Managed 12 member cabinet
-Organized student lobbying efforts at Iowa State House
-Assisted student body president with programming execution
College of Agriculture Senator, Government of the Student Body, Iowa State, 2000-2001
Conference Director, Big XII Student Government Leadership Conference, 2000-2001
President, Sigma Alpha Professional Sorority, 2001

Citizenship: U.S. and U.K.

Software: STATA, R, Matlab, ArcGIS, Geoda, PySAL, SAS, SPSS, GAMS, @Risk, Microsoft Office