The Pennsylvania State University

The Graduate School

Department of Agricultural Economics and Rural Sociology

SELF-EMPLOYMENT INCOME AND U.S. MIGRATION NETWORKS:

IS THERE A RELATIONSHIP?

A Thesis in

Agricultural, Environmental, and Regional Economics and Demography

by

Lucas Schaeffer

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2011

The thesis of Lucas Schaeffer was reviewed and approved* by the following:

Stephan J. Goetz
Professor of Agricultural, Environmental, and Regional Economics
Thesis Co-Advisor

Leif Jensen
Professor of Rural Sociology and Demography
Thesis Co-Advisor

Theodore R. Alter
Professor of Agricultural, Environmental, and Regional Economics

Ann Tickamyer
Professor of Rural Sociology
Head of the Department of Agricultural Economics and Rural Sociology

Edward C. Jaenicke
Associate Professor of Agricultural Economics
Coordinator, Graduate Program in Agricultural, Environmental, and Regional Economics

*Signatures are on file in the Graduate School

ABSTRACT

Self-employment has been an important driver of economic growth throughout the United States in preceding decades. Even though there are a greater number of self-employed within the economy, it has become relatively less profitable to own a business when compared to paid employment. Domestic migration, or the movement of population from one county to another within the United States, is another strong force that is shaping population change and regional economic development. Researchers have examined how domestic migration influences economic growth; however, no one has examined how it may relate to self-employment income success within the U.S. I draw upon inferences from migration, entrepreneurship, and social network theories to develop a model to test the relationship between domestic migration and self-employment income growth. Using migration flow data from 3,047 U.S. counties for the years 1995-2000, I find that high volumes of migration per capita from a county are not related to self-employment income growth from 2000 to 2007, although there is an inverse relationship between large inflows and growth. When including a variable that measures the diversity of a county's migration network (entropy), I find that the expansiveness of a network is positively associated with self-employment returns, especially within the most rural of counties. The diversity of migration networks is a more important determinant of entrepreneurs' success throughout the U.S. when compared strictly to the volume of migration flows. These findings have important implications for regional economic development, especially for rural areas that are experiencing large outmigration flows. If the migration network flowing out of or into rural areas is diverse, the self-employed are more likely to earn more than if the migration network is more homogeneous.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Stephan Goetz, my AEREC faculty advisor and mentor, for his strong guidance during the writing of this thesis. His insights and expertise in entrepreneurship have contributed greatly to my research. I would also like to thank Dr. Leif Jensen, who agreed to serve as a thesis co-chair. His knowledge of demographic theory and migration were also helpful to the success of this research. I would also like to thank Dr. Ted Alter for agreeing to serve as a member of my thesis committee. Dr. Brady Deaton was also incredibly helpful in the development and understanding of the migration entropy variable. Special thanks are due to Dr. Ted Jaenicke, who has assisted me through the graduate school process and ensured a meaningful experience at Penn State.

Finally, I would like to thank my family, friends, Dr. Hallie Savage, and Dr. Rod Raehsler. Each has proved to be a source of guidance, wisdom, motivation, and support throughout my academic career.

Chapter 1
Introduction

Entrepreneurs, or the self-employed, are becoming increasingly important in current economic times[1]. In the last 40 years, there has been large growth in the number of self-employed in relation to wage and salary workers, but at the same time entrepreneurial profits have been steadily declining. Even though individuals are becoming self-employed either by choice or force, they are failing to garner as much income as they could through being employed by others. The vitality of this growing class is important to the overall success of the economy. If profits of the self-employed continue to decrease in forthcoming years, the potential negative impact on consumer spending, government social programs, and the strength of regional economies will grow continuously stronger.

There has been a major push in regional economic development policy to emphasize "home-grown economic activity," or growth from small businesses within the home community, as a means of driving regional economies (Deller & Goetz, 2009). The tenets of the small business focus emerged from the works of Birch (1979; 1981), who found that growth within the economy was more attributable to small firms than large ones. Many times small business growth comes from entrepreneurs and those that they employ within a community. Even an individual entrepreneur can be considered a firm (Strauss, 1944). The potential problem associated with the push for home-grown activity in the form of self-employment is that if returns are lower than for paid employment, there may be incentives for individuals to switch from higher paying to lower paying work that contributes less to the local economy.

---

[1] Following past research by Low (2004), I use the terms "self-employment" and "entrepreneurship" interchangeably throughout this paper, although some researchers argue that there are important differences in how each conducts and builds business.

Returns to Self-employment

As stated previously, returns to self-employment within the United States are decreasing compared to paid employment[2]. Figure 1 shows the long-run trend in returns to self-employment in comparison to returns to paid employment. A value of one illustrates parity in returns, with values below one meaning that self-employment returns are lower than what is earned within paid employment. The only year in which self-employment returns were greater than paid returns was 1969. Since that time, returns have been variable, but the trend shows a generally decreasing ratio. During the 1970s, there was a rapid decrease in self-employment returns, with a rebound in the 1990s. This rebound never again reached 1970s levels, and by 2008 the self-employed earned less in relation to paid employees than ever before. They earned just $63 for every $100 earned in paid employment, or over 1/3 less.

Figure 1: Historical Self-employment Returns

[2] The Bureau of Economic Analysis (BEA) measures self-employment returns by the average income to nonfarm self-proprietors as reported by individual income tax returns. The return to wage and salary employment is the average yearly income of a worker employed by someone else.

Table 1 provides more detailed data for the years 2000-2007, the time period examined within this analysis. The average annual change in self-employment profits varied throughout the seven-year timespan, but decreased between years three times (2002-03, 2004-05, 2006-07). The decrease between 2006 and 2007 was drastic, and by 2007 returns to self-employment were actually lower than what they were in 2000. The ratio of self-employment to paid employment earnings also decreased rapidly throughout the first decade of the 21st century. In 2001, the self-employed earned nearly 90% of what wage and salary workers earned. In other words, for every $100 that a salaried worker earned, a self-employed worker earned $90. By 2007 the proportion had decreased to under 69%, a significant decline. At regional levels, there is also significant variation among states. In Oklahoma, the self-employed earned only 13% less than salaried workers, while in Florida, they earned 46% less (BEA REIS, 2010). The overall trend, however, also holds regionally: self-employment is becoming less attractive from a strictly monetary viewpoint.

In nonmetropolitan (nonmetro) areas, the self-employed earn proportionately less than paid employees when compared to metropolitan (metro) areas. Table 1also provides self-employment income figures for nonmetro areas for 2000-2007. At the beginning of the period, entrepreneurs in nonmetro areas earned just 75% of what paid employees made. By 2007, the proportion had dropped to 61%, a dramatic decline. Throughout the decade, nonmetro self-employed earned less than metro self-employed. This is disconcerting, as it illustrates the fact that entrepreneurs in rural areas are struggling to create businesses that contribute high value to the local economy.

Table 1: Self-employment Returns, 2000-07

Year	Self-employment Returns	Ratio of Returns vs. Paid Employment Returns	Average Annual Change
U.S. Total			
2000	$30,851	0.880	-
2001	$32,318	0.897	4.76%
2002	$32,577	0.889	0.80%
2003	$31,931	0.846	-1.98%
2004	$33,312	0.846	4.32%
2005	$32,964	0.808	-1.04%
2006	$34,082	0.798	3.39%
2007	$30,671	0.688	-10.01%
Nonmetropolitan Counties			
2000	$18,819	0.749	-
2001	$20,132	0.775	6.98%
2002	$19,983	0.746	-0.74%
2003	$19,949	0.721	-0.17%
2004	$20,858	0.725	4.56%
2005	$20,877	0.703	0.09%
2006	$21,390	0.691	2.46%
2007	$19,713	0.611	-7.84%

Source: BEA REIS, 2010.

The disparity between self-employment and paid employment returns may not be captured fully with these data, as fringe benefits for paid employment such as contributions to employee pension funds, insurance payments, and payments to government social insurance are not factored into earnings. If so, it is likely that the income gap between the two types of employment would be greater. Conversely, some argue that a portion of the difference in earnings is also a function of underreporting of income by the self-employed due to tax incentives (Hamilton, 2000; Taylor, 1996; Blau, 1987; Rees & Shah, 1986). Higher underreporting would also serve to widen the gap in pay between the two types of employment. The magnitude of the negative underreporting effect in comparison to the positive fringe benefit effect influences how large the difference in earnings is. For the data included within this study, the BEA adjusts self-employment earnings to incorporate underreporting. This aids in showing the true gap that exists between entrepreneurial and paid returns.

If the self-employed earn so much less than the paid employment workforce, why do so many participate in the entrepreneurial economy? One explanation is that individuals are forced into self-employment because of the lack of employment opportunities in the wage and salary sector (see Parker, 1996 for a brief explanation of the "recession push" theory). A second well-documented explanation is that self-employment returns are also nonmonetary in nature. For example, entrepreneurs value and gain utility from having greater work freedom. It is not possible to quantify the nonmonetary returns to entrepreneurship with the data being used within this analysis; however, a discussion of why they factor into returns is important and located in Chapter 4.

Growth in Self-employment Numbers

Even though relative monetary earnings for the self-employed have decreased, growth in the number of self-employed has been a major driving force in the development of new jobs throughout the United States[3]. Figure 2 illustrates the growing importance of self-employment. In 1970, self-employment accounted for around 10% of total full and part-time employment. By 2008, the proportion of self-employed had increased to 20%. In other words, one-fifth of the U.S. workforce was involved in entrepreneurial activity by the late part of last decade.

[3] Self-employment is measured by the BEA as the number of nonfarm self-proprietors.

Figure 2: Self-employment as a Percent of Total Employment

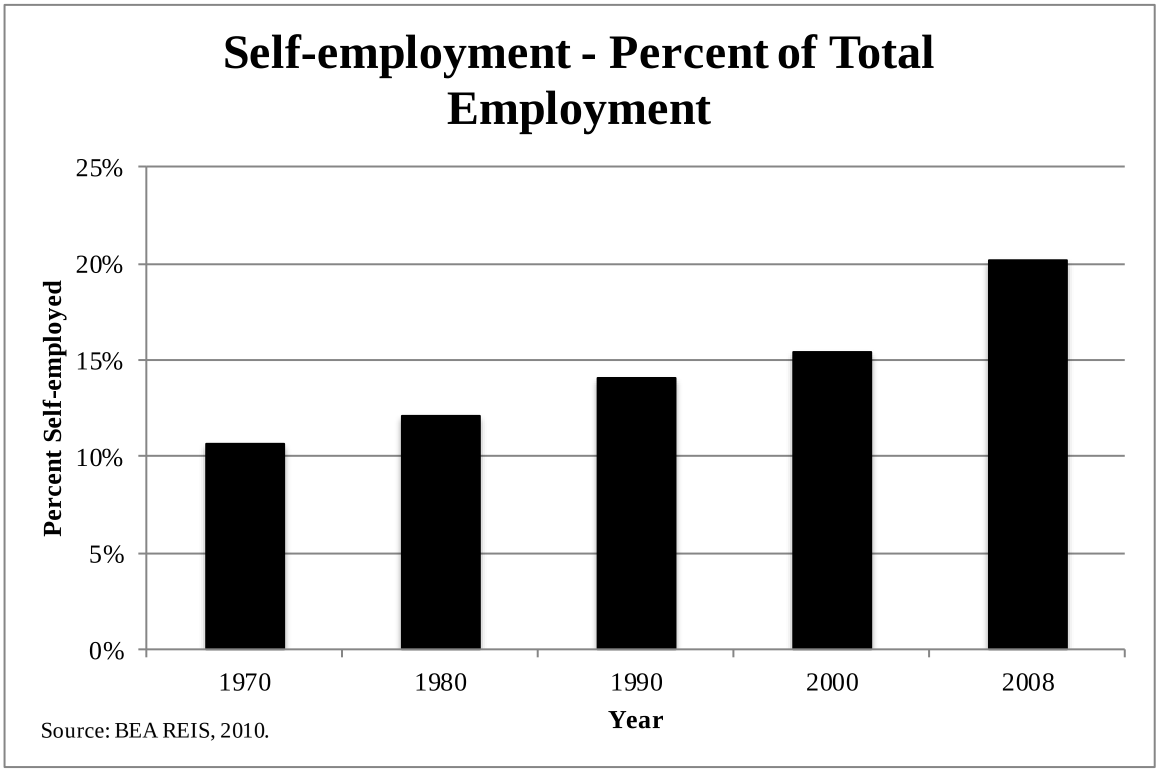


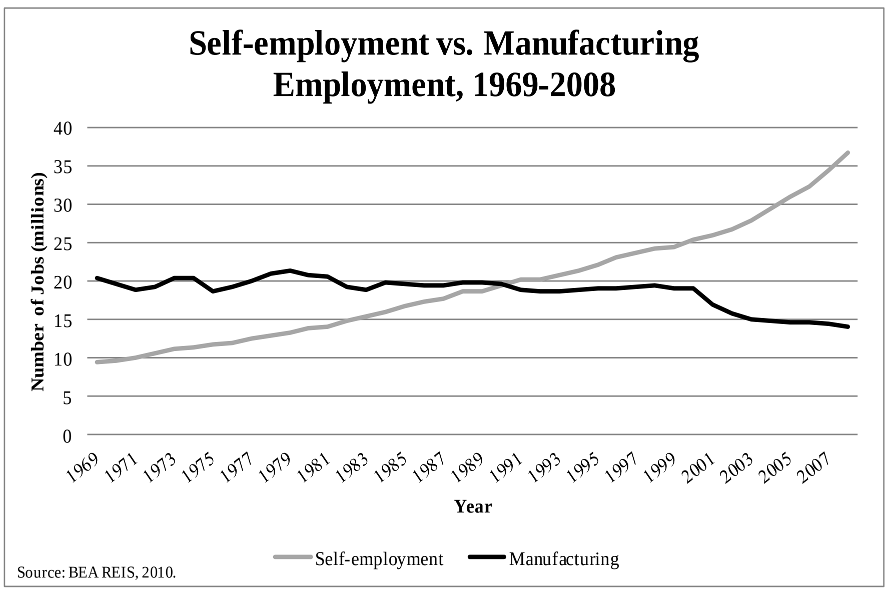
Table 2 shows the increasing trend in self-employment from 2000 to 2007, the time period of interest within this study. Self-employment growth occurred more rapidly during this decade than any other. The average annual increase in number of self-proprietorships was 4.7%, more rapid than any employment sector of the economy other than real estate (6.1%) and mining (5.6%) (BEA REIS, 2010). Nonmetro areas have a higher proportion (22%) of self-employed in the total workforce. Combined with the fact that rural entrepreneurs earn only two-thirds of what paid employees do, a higher proportion of self-employed may lead to declining local income for nonmetro areas.

Table 2: Growth in the Number of Self-employed

Growth in Self-employment, 2000 to 2007		
Year	**Self-employment**	**Annual Percent Increase**
2000	25,536,800	-
2001	25,998,200	1.81%
2002	26,762,100	2.94%
2003	28,001,500	4.63%
2004	29,541,700	5.50%
2005	31,122,400	5.35%
2006	32,381,600	4.05%
2007	34,459,700	6.42%
Source: BEA REIS, 2010.		

Some growth in self-employment numbers may be related to the decline in U.S. manufacturing employment. Figure 3 illustrates the importance of the rise in self-employment in a time of declining manufacturing jobs. Self-employment has risen from under 10 million jobs in 1969 to over 35 million in 2008, nearly a quadrupling in the number of self-employed. During the same time period, manufacturing has lost a total of over 5 million jobs, or over 30% of the total workforce of 1969. While it is important to note that some manufacturing jobs were "lost" due to the switch from the Standard Industrial Classification (SIC) system to the North American Industrial Classification System (NAICS) between 1999 and 2000, the overall declining trend is still clear: individuals may increasingly be turning to self-employment at least partly due to job

Figure 3: Self-employment and Manufacturing Employment Comparison



loss in other sectors of the economy. It is easy to see why self-employment has been a major focus of research during the preceding few decades. Without this growing class of individuals, local economies are likely to be much worse off.

The rise in self-employment has important implications for the economic and employment growth of regions. Using panel data from 23 OECD countries, Thurik et al. (2008) find that higher entrepreneurial activity reduces unemployment in subsequent periods. Data from the United States alone also show that entrepreneurship, especially within the services sector, is important to employment growth (Bednarzik, 2000). New firms formations, another form of entrepreneurial activity, are positively associated with growth in regional employment (Acs & Armington, 2006). The empirical consistency of this "entrepreneurial" effect (Audretsch & Thurik, 2000), or the increase in employment due to new firm start-ups, has been debated by many researchers (see Thurik et al., 2008 for a discussion of the literature). Nevertheless, the number of self-employed has increased so rapidly over the previous four decades that it is clear that this sector of employment is important to the economic vitality of the U.S.

Layout of Thesis

While the trend of decreasing entrepreneurial returns coupled with an increasing number of self-employed has important economic implications, there have been very few studies examining how returns to self-employment may be affected by other economic, demographic, or regional trends. This paper explores the potential linkages between returns to self-employment and one particularly strong demographic trend: domestic migration within the United States. More specifically, I examine the empirical relationship between profits and the characteristics of migration networks. The presence of migration networks has the potential to influence self-employment profits in both origin and destination locations by way of increased (decreased) levels of human capital and high (low) diversity in knowledge and information flows. To measure the extensiveness of migration networks, I use the number of inmigrants or outmigrants per capita (volume), along with a newly applied migration entropy variable that measures the dispersion and diversity of migration networks. Section II of this paper goes on to describe internal migration in the contexts of the United States, Section III relates U.S. migration flows to social network theory, Section IV describes the data and methods utilized for the analysis, and Section V provides a discussion of the results and concluding remarks.

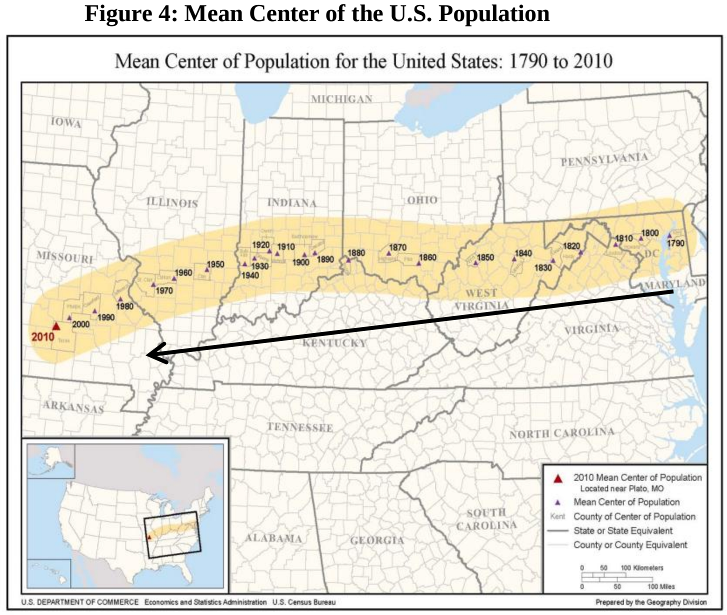
Chapter 2
U.S. Internal Migration

Domestic migration has received less attention in recent years in the literature, but it is still one of the most important demographic processes for the U.S. in terms of population growth and redistribution. Rates of natural increase - the number of deaths in the population subtracted from the number of births – are relatively low across the United States, with the majority of population increase or decline determined by migratory patterns of the population. Knowledge about the causes and consequences of internal migration is still somewhat cryptic (Greenwood, 1975), but there are some historically consistent characteristics that describe domestic migration flows.

Regional and Rural Characteristics of Domestic Migration Streams

Throughout the last 50 years, internal movement of population has been a major force in determining the distribution of the U.S. population. Figure 4 shows the mean center of population of the United States since 1790. The mean center represents the point at which a flat, featureless map of the U.S. would balance if identical weights were placed where each person resided during the latest Census year (U.S. Census Bureau, 2001).

Figure 4: Mean Center of the U.S. Population



The figure shows a clear picture: the mean center of the population is steadily moving towards the South and West, which is mostly a function of domestic migration to these regions from the Northeast and Midwest.

Table 3 provides more concise historical data on net migration for the four Census regions of the United States. The Northeast and Midwest regions have continuously lost population to the South and West regions. The Northeast and Midwest have lost 4.6 million and 3.2 million people, respectively, through outmigration since 1965. The South has gained the most population through inmigration, with an increase of over 5.6 million people. Migration during the 1995-2000 time period used within this study follows the same trend as the historical perspective. The Northeast and Midwest regions lost population, while the South and West gained (Franklin, 2003a). The West, however, had a much lower net migration rate than in previous years, signaling a slowdown in migration to that region. The South, on the other hand, continued to be the most attractive for migrants, gaining nearly 1.8 million people. This population redistribution through internal migration has important implications for regional economic development.

Table 3: Historical Regional Migration Patterns

Migration by Region of the United States

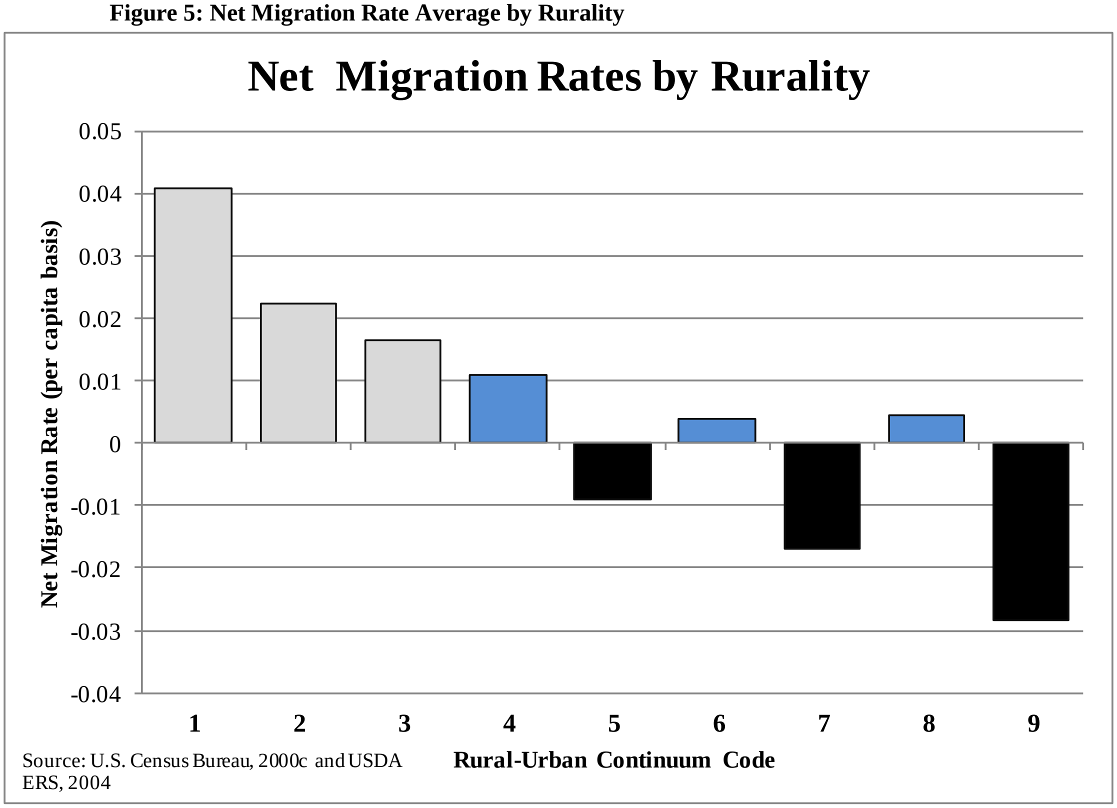
	1965-1970			1975-1980			1985-1990			1995-2000		
	In	*Out*	*Net*	*In*	*Out*	*Net*	*In*	*Out*	*Net*	*In*	*Out*	*Net*
Northeast	1,273	1,988	-715	1,106	2,592	-1,486	1,604	2,720	-1,116	1,537	2,808	-1,271
Midwest	2,024	2,661	-637	1,993	3,166	-1,173	2,324	3,172	-848	2,410	2,951	-541
South	3,142	2,486	656	4,204	2,440	1,764	4,769	3,344	1,426	5,042	3,243	1,799
West	2,309	1,613	696	2,838	1,945	893	2,827	2,289	538	2,666	2,654	12

Source: U.S. Census Bureau Census figures from 1970, 1980, 1990, and 2000.

Since the 1970s, U.S. migration patterns between metro and nonmetro areas have been more dynamic and changing in nature. Urbanization by way of migration dominated between the Great Depression and the 1970s "rural renaissance," where a reversal in nonmetro outmigration led to more rapid population growth in rural areas (Beyers & Nelson, 2000). The rural rebirth of the 1970s waned in the 1980s, but nonmetro areas again saw significant numerical gains from metro areas during the 1995-2000 time period. There were over 6.1 million migrants from metro to nonmetro areas and 5.6 million migrants from nonmetro to metro areas, leading to positive absolute net migration for nonmetro areas (Schacter, Franklin, & Perry, 2003).

Net migration *rates per capita*, instead of actual counts, show that rural areas overall are losing migrants compared to the total population. Figure 5 shows net migration rates for metro and nonmetro areas based on the United States Department of Agriculture (USDA) Economic Research Service's (ERS) Rural-Urban Continuum Code. The code indexes the rurality of a county based upon certain population characteristics and the proximity to metro areas. Counties with a score of 1 to 3 are metro counties, with those from 4-9 considered nonmetro. A full description of each code can be found in Appendix A.

Figure 5: Net Migration Rate Average by Rurality



Bars of the graph in gray represent metro counties. The average net migration rate across metro counties is positive, and decreases with higher rurality. Nonmetro net migration rates are the most interesting. Blue bars represent nonmetro counties that are adjacent to metro areas, and black bars represent nonmetro counties that are not adjacent to metro areas. Counties that are adjacent to metro centers all have small, but positive net migration rates (a net migration rate of .01 means that a county's population increased by 1% due to migration). The most rural of counties, or those that are not adjacent, all experience negative net migration rates that decrease with subsequently higher levels of rurality. There is a clear distinction between the two groups of nonmetro counties. Those which are located close to urban centers are more attractive to migrants, possibly due to the proximity to services that urban areas can provide such as healthcare, employment, and infrastructure (Kilkenny & Johnston, 2007). The most rural of counties are losing population, which has potentially severe negative consequences for the economic vitality of those regions.

Education Levels of Migrants and Implications for Origin Communities

Another important dynamic of internal migration streams is the education of migrants. Migration theory states that migrants are positively selected – that they have higher levels of education than their peers who do not migrate. Sjaastad (1962) was one of the first to assess the effects of human capital on migration propensity, finding that those with a higher level of education are more likely to move. Certain employment opportunities with advantages such as rapid promotion and higher wages are available only to those with higher levels of human capital, so the educated have more choices in employment, both near and far. Additional information about job opportunities located in distant lands is only available to the more educated, again providing them more enticing employment options in different geographic areas (Schwartz, 1976). Individuals who have high skill levels are also likely to migrate to where they can best make use of those skills (Borjas, Bronars, and Trejo, 1992; Nelson, 1959). Migration, for individuals, is thus an investment in earning higher future incomes (Bowles, 1970). In other words, those with higher human capital migrate to where they can earn the highest return on their investment in education.

Data from U.S. Census Bureau's Current Population Survey (CPS) show that the propensity to migrate increases with higher levels of education. Table 4 shows the absolute number of movers as a percent of the population with the same education level. Over the years 1997-2000, the total percent of the population moving in each year was around 5%. The percent of the population that migrated was less than the overall average for lower levels of education, with those in the lowest educated category migrating the least. In every year, increasing levels of education, up to a bachelor's degree, led to a higher propensity to migrate. These data agree with historical theoretical migration frameworks: the most educated are also the most mobile.

Table 4: Education of Migrants, Evidence from the CPS

	1997-1998		**1998-1999**		**1999-2000**	
Education Level	***Movers***	***Percent of Pop.***	***Movers***	***Percent of Pop.***	***Movers***	***Percent of Pop.***
All Levels	*8,136*	*4.7%*	*9,113*	*5.2%*	*9,698*	*5.5%*
Less than 9th grade	281	2.2%	324	2.6%	493	4.0%
Grades 9-12, no diploma	613	3.7%	652	4.0%	759	4.8%
High school graduate	2,346	4.0%	2,800	4.8%	2,868	4.9%
Some college or AA degree	2,226	5.2%	2,397	5.6%	2,520	5.7%
Bachelor's degree	1,843	6.5%	1,984	6.7%	2,141	7.2%
Prof. or graduate degree	827	6.1%	955	6.7%	921	6.1%

Source: U.S. Census Bureau CPS, 2010.

Data on domestic migration from the 2000 Census coincide with that from the CPS. When looking at migration of just 25-39 year olds, it is clear that the most educated are the most likely to migrate (Franklin, 2003b). Table 5 provides data on the education of migrants in this age group. The college educated have a much higher propensity to migrate than their non-college educated counterparts: 37% of the college educated population were movers in comparison to just 22% of the non-college educated population. Migration of the non-college educated is dominated by within state movers (12.5% of the total non-college educated moved to a different county in the same state). Just over 9% of the population lacking a college education moved to a new state. Conversely, migration of the college educated is dominated by movers to a new state. Over 20% of the college educated population moved to different states, while same state movers comprise only 17.5% of the total population. The data confirm the theory that the average distance of migration increases with higher levels of human capital because of higher potential returns in relation to the costs of migration (Sjaastad, 1962; Courchene, 1970; Schwartz, 1973; Folger & Nam, 1967).

Table 5: Education of Migrants, Evidence from Census 2000

Data on Education of Migrants from the Ages 25 to 39							
	Total Population	**Within State Movers**	**% of Total Population**	**Different State Movers**	**% of Total Population**	**Total Movers**	**% of Total Population**
Total	62,660,694	8,697,378	13.88%	7,759,683	12.38%	16,457,061	26.26%
College Educated							
Metro	15,188,587	2,589,921	17.05%	3,073,522	20.24%	5,663,443	37.29%
Nonmetro	1,688,718	368,129	21.80%	315,333	18.67%	683,462	40.47%
Total	16,877,305	2,958,050	17.53%	3,388,855	20.08%	6,346,905	37.61%
Not College Educated							
Metro	36,588,662	4,335,012	11.85%	3,487,651	9.53%	7,822,663	21.38%
Nonmetro	9,194,727	1,404,316	15.27%	883,177	9.61%	2,287,493	24.88%
Total	45,783,389	5,739,328	12.54%	4,370,828	9.55%	10,110,156	22.08%
Source: Franklin, 2003b and Author's Calculations							

The metro/nonmetro breakdown of population migration by education also has important implications for rural areas. Of the non-college educated, the nonmetro population is more likely to move than the metro population, and when those residents do move they tend to move to other counties within the same state. What is most important for rural areas, however, is that the propensity for the highly educated to leave is much higher than the non-educated. 40.5% of the rural college-educated population outmigrates from their origin counties as compared to just 25% of the non-college educated. The nonmetro educated population is also more likely to migrate than the metro college-educated (only 37% of the metro educated population are movers). While there is no distinction as to where they are migrating to, it is likely that the educated rural population is moving to more urbanized areas where individuals can garner higher returns to their investments in human capital (Kusmin, Gibbs, & Parker, 2008; Goetz & Rupasingha, 2003). Outmigration of the young, educated class from rural areas constitutes the "rural brain-drain," which has potentially negative implications for economic growth in rural areas (Artz, 2003).

Rural brain drain is a topic of great concern in the regional economic development sphere. Rural outmigration counties do not all suffer from economic stagnation and high rates of poverty (McGranahan, Cromartie, & Wojan, 2010; Johnson, 2011). Net outmigration counties with high levels of natural amenities (i.e. mountains, lakes, many days of sunshine) are likely to be less susceptible to economic hardship (for review see Irwin et al., 2010). Some net outmigration counties prosper, but others without certain characteristics struggle. Those counties with high outmigration rates combined with lower levels of human capital (resulting from high brain drain) are much more likely to suffer persistent poverty (McGranahan, Cromartie, & Wojan, 2010; Findeis & Jensen, 1998). Artz (2003) finds that the most rural of counties are the most likely to suffer from high rates of brain drain. Using data from the CPS for the years 1989-2004, Domina (2006) also finds that education is in fact the most important predictor of outmigration from nonmetro areas, which coincides with early research in the migration field (Wolpert, 1965). Continued high levels of brain drain from rural areas leads to depletion of human capital, reinforcing high poverty (Schacter, Jensen, & Cornwell, 1998). If places with increased levels of human capital enjoy more rapid growth and higher productivity, as past research suggests (see Glaeser, 2001; Florida, 2002b; Lobo & Smole, 2002; Rauch, 1993), nonmetro areas with high outmigration rates will continue to struggle to find economic prosperity in the future (Domina, 2006).

Migration's Potential Impact on Self-employment Success

Outmigration of the educated from rural areas may be related to entrepreneurial success. Drabenstott (2001, p. 13) cites brain drain as one of the most significant issues that impedes "stoking entrepreneurialism" and creating high-value opportunities for the self-employed in rural areas. Entrepreneurs have higher levels of educational attainment in comparison to paid employees. Those with higher education are likely to be the ones with the ability to recognize new and profitable business opportunities (Acs & Armington, 2006). Conversely, those with lower human capital accumulation are likely to be those who fail in business venturing. Robinson and Sexton (1994) find that extra years of schooling lead to greater entrepreneurial success in the form of progressively higher self-employment earnings. Hamilton (2000) finds that progressively higher schooling, from a high school dropout to a college graduate, leads to higher hourly wage returns in self-employment. Being a high school dropout negatively influences self-employment returns, while possessing a college education raises the productivity of entrepreneurs (Hamilton, 2000). Returns to self-employment for the college educated holds across racial categories, as both Asian and African Americans who have a college degree and are self-employed earn a great deal more than their less educated counterparts (Bates, 1997).

The most important common characteristic between migrants and entrepreneurs shown here is a higher level of human capital. If a high proportion of highly educated are leaving rural areas, who is left within the county to create entrepreneurial success? Is the lower educated population that remains capable of creating economic growth through successful entrepreneurship? Concurrently, the same counties that are losing their educated also suffer from high poverty, limiting the potential market for an entrepreneur to sell products or services. Based upon these insights, high levels of outmigration from rural areas should negatively affect self-employment income in the county through the reduced ability of the remaining population to create successful and profitable entrepreneurial ventures.

The existence of the migration network as a potential knowledge transfer pathway must also be taken into account. Even though migrants are taking high levels of human capital with them, they may also serve as conduits through which information flows to home counties. The negative impact of outmigration of the educated may be counteracted by increased transfer of novel knowledge. A discussion of how migration networks may influence success of entrepreneurs follows in the next section.

Chapter 3
Network Theory

Social Network Theory and Application to Entrepreneurship

The study of social networks is becoming an important research area within many disciplines. Sociologists, anthropologists, economists, and political scientists have begun to use the tools of social network analysis to analyze the systematic relationships that exist between social actors (Borgatti et al., 2009). The implications of these relationships are important for answering questions about social, economic, and political structural environments (Wasserman & Faust, 1994). To frame the concept of a social network in basic terms, it is a set of actors, nodes, or agents that may have relationships to one another (Hanneman & Riddle, 2005). In fact, "relations defined as linkages among units are a fundamental concept of network theories" (Wasserman & Faust, 1994, p. 4). These relationships may be economic, political, or personal in nature, but each type may be systematically analyzed using social network analysis. Network analysis is rooted in mathematical and statistical foundations; however, the ideas behind the central concepts are motivated by social theory (Wasserman & Faust, 1994). The combination of grounding in both empirical and theoretical bases makes network analysis a strong analytical tool. In the past, researchers have used network analysis at both the macro and micro levels to study phenomena as widespread as a country's importance in world economic systems (Snyder & Kick, 1979) to medical innovation spread (Coleman, Katz, & Menzel, 1966).

In this paragraph, I provide a brief description of some of the important concepts in network theory. Much of the explanation of these ideas follows closely to that of Borgatti et al. (2009), Tutzauer (2007), and Wasserman and Faust (1994). Each of the actors within the network is termed a *node*. Nodes can be individuals, corporations, or collective units. Within my research, the node is measured as a collective unit (county). Each node has a number of *relational ties* to one or all of the other nodes contained within the network. An important characteristic of a tie is that it establishes a definite linkage between a pair of nodes. The relational tie within my research is population movement between two locations. A migrant moving from one county to another establishes a linkage between these two counties. The strength of this linkage is partially determined by the number of migrants moving between the two locations. Using the existence or strength of a tie, the power of the node's location within the network can be assessed using various *centrality* measures.

Different centrality measures lead to varying measures of prominence within the network. Degree centrality, originally formulated by Freeman (1979), is one example of a centrality measure. Degree centrality counts the number of ties that a node has, in effect cataloging a node's level of activity. If the ties are bi-directional (with numerical value of flows differing depending on flow to or from a node), both *indegrees* and *outdegrees* can be measured. For the case of migration networks, indegrees are the number of migrants moving to a county, and outdegrees are the number of migrants leaving a county. Other centrality measures include *betweenness* and *closeness,* which show the extent of a node's control and communication efficiency. A relatively new centrality value termed *entropy* provides insights into the dispersion of network ties among other nodes in the network (this will be discussed fully in subsequent text). The node's *outcome*, or success in some activity, normally relates to its centrality within the network. High values of centrality measures often lead to better outcomes. The consequences of the network on some activity depends on the fundamental theory that information passes along the ties within the network, and those nodes in advantageous, prominent locations within the network will control and receive the most diverse information. In this research, I examine the consequences of network centrality, as measured by entropy, on self-employment income growth.

In the field of regional economic development, social network analysis is emerging as a new field of study. Researchers are beginning to realize that the existence of interpersonal networks has important implications for economic development at the regional level, specifically within entrepreneurial development (Fortunato & Alter, 2011). There is a wide breadth of literature relating self-employment success and access to social ties and networks (see Hoang & Antoncic, 2003, for an extensive review), but none provides empirical analysis of entrepreneurial success in relation to migration networks specifically at the macro-level. Aldrich and Zimmer (1986) were the first to relate entrepreneurial success to structured social networks. The authors place entrepreneurship within the confines of Granovetter's (1983) embeddedness theory, where business ventures are either facilitated or confined by social ties and networks. The most pertinent theoretical postulation that Aldrich and Zimmer make in the contexts of this research study is that exposing entrepreneurs to a wider diversity of social ties will heighten the opportunities available to them, increasing the probability of success. Dubini and Aldrich (1991) expand upon the above theoretical framework, but still note the importance of diversity within networks for entrepreneurial success. Using data on German business start-ups, Brüdler and Preisendörfer (1998) empirically test if the diversity of networks (measured by a proxy for Granovetter's weak ties) is important in entrepreneurial growth and success. They find a significant, positive relationship between the two.

In networks with a large number of weak ties, or those characterized by "hubs," nodes are separated from all others by a small number of steps (Barabási, 2002). Information about new ideas is more easily propagated in networks with a large number of weak ties, or scale-free networks (Ogle, 2007). U.S. migration networks have this characteristic: certain counties are hubs of migration, and they serve as connections through which greater information may flow between many of the smaller county networks[4].

Overall, diversity in information sources is potentially more important than the actual number or volume of flows. Page (2007) insists that diverse information flows are more conductive for problem solving or generating novel ideas. Eagle, Macy, and Claxton (2010, p. 1029) articulate Page's ideas well when speaking about social network analysis: "highly clustered, or insular, social ties are predicted to limit access to social and economic prospects from outside the social group, whereas heterogeneous social ties may generate these opportunities from a wide range of diverse contacts." Furthermore, the researchers find a relationship between diverse communication networks and regional economic success in the UK. For higher economic prosperity, diversity, and *not* absolute volume, was the most important characteristic of the network analyzed. Within the specific context of entrepreneurship, diverse information flows are highly beneficial to successful business ventures (Hoang & Antoncic, 2003). Applying these results to migration networks tells us that diverse and high-value information flows should arise through heterogeneous, scale-free systems instead of simply high volume homogeneous flows.

Migration Networks as Sources of Information

While some have looked at migration streams as networks and conduits of information, no one has examined how specific measures of centrality as defined by social networking analysis are related to regional economic development in the U.S.[5] It is beneficial to explicitly define what a migration network is in the contexts of social network analysis. Haug (2008, p. 588)

[4] An analysis of the distribution of outmigrants and inmigrants across counties shows that the U.S. migration network follows the properties of a scale-free network.

[5] At the international level, there is only one study known to the author that specifically examines the relationship between migration network centrality and economic factors. Nogle (1994) analyzes intra-European Union migration streams, finding certain cliques of countries whose migration streams are tightly wound. The researcher also attempts to determine the influences of macro factors on migration.

provides a concise explanation of a migration network: “a migration network can be defined as the composite of interpersonal relations in which migrants interact with their family and friends,” and the subsequent social network formed by these interactions “provides a foundation for the dissemination of information as well as for patronage or assistance.” In the context of this study, individual interpersonal relations of migrants are assumed to increase proportionately with aggregate size of the migration network of the county. Micro-level interactions and relationships are aggregated to form a comprehensive macro-level measure of access to information and knowledge for both origin and destination counties by way of migrant flows.

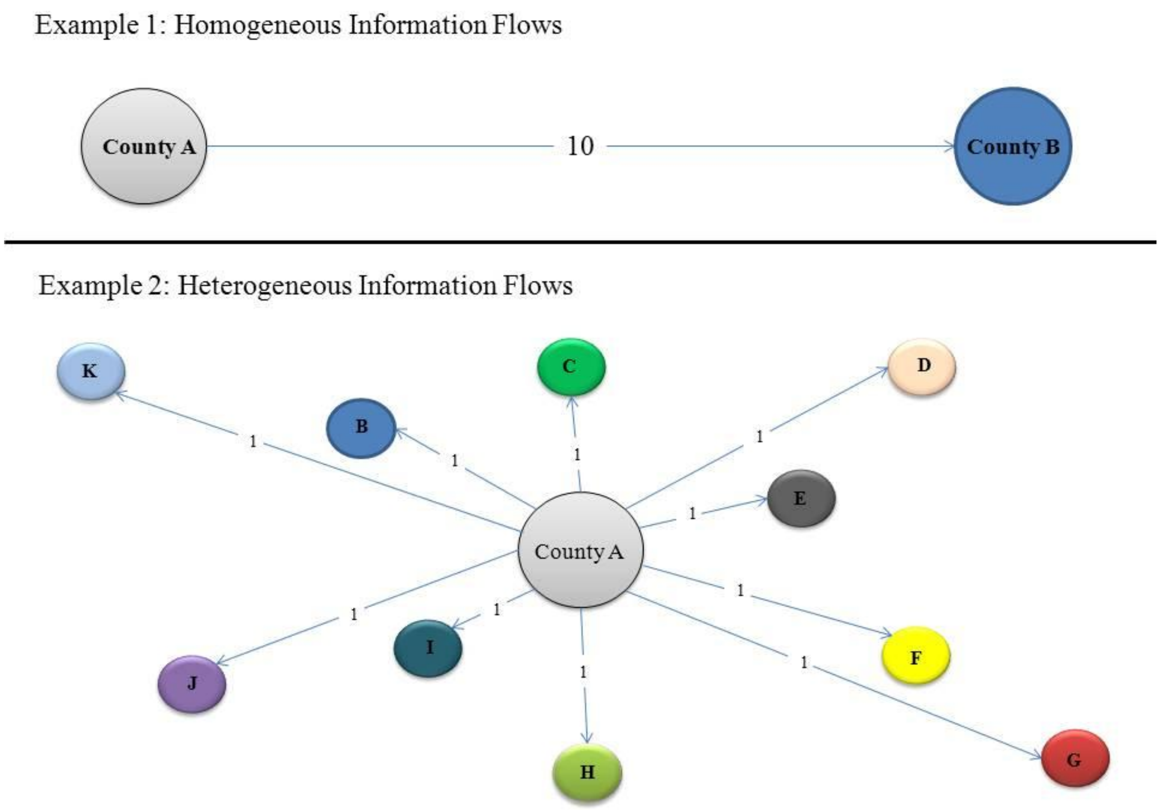
In international migration literature, the importance of migration networks as sources of information for home regions is well established. Migrants retain connections in their home community. Wellman (1983, p. 164) states that “migration is rarely a once-and-for-all, uprooting, and isolating experience. Rather, migrants travel and communicate back and forth between their residence and ancestral homeland.” Here, the ancestral homeland is the origin county instead of a different country, but the basic theoretical idea of retained community ties remains true.

Instead of viewing outmigration of the highly educated as a completely negative phenomenon for the home community, the network created by those leaving should be viewed as “an asset that can be mobilized” (Meyer, 2001, p. 97). Even though they are currently residing in a different area, migrants still have a vested interest in their home community. Home communities may benefit from the embedded knowledge and extensive networks that have been built by outmigrants (Meyer, 2001). This is sometimes termed “brain circulation”, or the flow of new ideas back to the home area by way of knowledge transfers from those who have left (Gaillard & Gaillard, 1998; Johnson & Regets, 1998). While this has not been systematically examined on the domestic level within the U.S., it is likely that brain circulation functions in much as the same way as it does internationally. Those who have left retain ties within their origin community and subsequently pass knowledge acquired in the new location to those at home. It is also important to note that decreases in the cost of transporting information (i.e. easier access to internet and cellular phones) over preceding decades has facilitated increased communication and knowledge sharing by migrants.

While there has been increased interest in how outmigration (rural brain drain) negatively affects economic development in rural America, there is a dearth of literature that examines specifically how migrant networks affect economic activity in origin and destination locations either negatively or positively. Millimet and Osang (2007) study how state-to-state internal migration networks affect trade and find that migrants are important facets for growth in both origin and destination. The researchers conclude that outmigrants bring new information to destination areas about products and services that were produced in the home location. Migrants also provide information about products and services produced in the destination area to those still within the home community. This novel information flow from migrant sources can be invaluable for entrepreneurs. The knowledge gained through the networks can inform individuals of what products and services are successful in other areas. If the same products are not being offered in the home destination, a nonmigrant may capitalize on the undersupply by beginning their own business. Just as importantly, knowledge of business failures also passes via migrants. This helps to inform potential entrepreneurs of ventures that have been unsuccessful in other areas. Again, nonmigrants would likely not have known about fruitful or failed entrepreneurial endeavors elsewhere if it were not for the presence of a migrant network to provide the evidence. Aside from knowledge of outside markets, information about best business practices may flow from migrant to nonmigrant. Improved business practices cut the costs of operating a self-owned business, increasing profitability and success of the firm.

For the sharing of new information, it may not be sufficient to simply have a large network of migrants. As discussed above, diversity in information flows is potentially more important than the actual volume of flows. If migration streams to and from a county are relatively homogeneous, there is less differentiation in the ideas and knowledge that is transmitted through the network, which influences the amount of novel information that is accessible to current and potential entrepreneurs. A stylized example is illustrated in Figure 6.

Figure 6: Stylized Example of Information Flow Diversity



This example shows two different hypothetical migration streams from County A. In the first, 10 people leave county A and all 10 settle in county B. In this example, there is limited breadth of information flowing to county A because all migrants are settled in the same location. Only knowledge gained in the one county will flow back to County A. If, on the other hand, the 10 migrants are split among counties B, C, D, E, F, G, H, I, J, and K, the likelihood of heterogeneous information flows back to the home county is more likely. Diverse information from the different destination counties provides a wider and more comprehensive knowledge flow back to County A. The same simulation would hold for inmigration. Counties that receive a high number of migrants from a homogeneous source will see less diversity in information flows than a county that receives the same number of migrants from a heterogeneous set of counties.

In the following chapter, I discuss the variables used to measure diversity and volume of migration flows. I hypothesize that having a high volume of migration from a county will negatively influence entrepreneurial success of that location, especially in nonmetro areas. This hypothesis is based on the fact that even though migrants are sources of knowledge for both origin and destination counties, they are also the most highly educated. Thus, when a county experiences a high level of brain drain, there will be a negative relationship between self-employment returns and outmigration volumes (even though the effect may be mitigated through knowledge transfers). When a county experiences a high volume of inmigration, I hypothesize that self-employment returns will be higher because those moving to an area are the more educated and able to recognize markets for potential entrepreneurship. Counties with high diversity in migration streams, as measured by the entropy centrality measure, should have improved self-employment income growth outcomes because of their preferential location in the network.

Chapter 4
Data and Methods

Model and Data Sources

To model growth in entrepreneurial success, I extend a model used by Goetz & Shrestha (2009). The authors use a Mincer-style growth equation to measure the growth in self-employment income. Income growth for county *i* from year *t* to *t+d* is a function of migration activity and networks of the county (M), initial levels of self-employment income (SE_i), a set of county and economic indicators (C), human capital and experience variables (H), and county demographic variables (D). The model is as follows:

$$Self - Employment\ Income_{i,t+dt} = f\left(M_{i,t}\ , SE_{i,t}\ , C_{i,t}\ , H_{i,t}\ , D_{i,t}\right)$$

where *t* is the base year 2000 or closest to the base year as possible and *d* = 7 indicates the length of the time period. Using the change in entrepreneurial depth between two time periods instead of from just one time period helps to avoid the issue of simultaneity[6], even though it is a more difficult model to estimate (Goetz & Shrestha, 2009).

Counties within the U.S. are the level of classification within this study. This research includes data on 3,047 county and county-equivalent places within the country. It is important to note that county-wide measures serve as a proxy for the average individual within the county[7]. By using a nationwide geographic basis, it is possible to gain a wider understanding of how

---

[6] It is nearly impossible for change in entrepreneurial depth over time to have a feedback effect on initial characteristics of the county and its population.

[7] It is important to note that ecological fallacy, or applying individual level characteristics to aggregate population levels, may be an issue. County level measures do serve as a proxy, however, for the characteristics of average potential entrepreneur and the economic conditions that they face. Using county level data also eliminates certain selection biases due to inclusion of the entire population's characteristics.

proprietor incomes are influenced by population movement at a large scale. Because of comparability across time and the changing nature of boundaries, I do not include the Alaskan census places. I also exclude counties within Hawaii because of the distinct geographic location of that state. Certain independent cities within Virginia present data issues as well. The BEA and the U.S. Census Bureau combine statistics from the independent cities with surrounding counties differently. Because of this, I do not include the independent cities that are incompatible across data sources.

Counties within Texas account for over 8 percent of the observations, with other large concentrations within Georgia, Kentucky, and Missouri. Not all geographically large states have many counties, and vice versa. The size of counties vary dramatically throughout the U.S., with larger counties concentrated in the less densely populated West, and smaller counties in the East. California is one of the largest states and accounts for over 4% of the total land area of the U.S.; however, it has a relatively small number of counties that represents only around 2% of the total observations. Kentucky, conversely, accounts for only 1% of the total U.S. landmass, but nearly 4% of all observations.

The first major data source of this research is the BEA Regional Economic Accounts (REIS). The BEA provides yearly information on regional and local economic activities. Employment and income data, broken down by SIC and NAICS code, are available back to 1969 for geographies down to county level. The REIS data are some of the most useful to regional economists because of their yearly availability and continuity over time.

The second major data source within this research is the U.S. Census Bureau’s 2000 Census. The Census provides a wealth of information on many demographic, housing, and economic characteristics of the U.S. population (U.S. Census Bureau, 2000c). The data on population characteristics provided by the Census come from nearly 19 million long-form surveys that were returned during 2000. This sample represents approximately 1 out of every 6 U.S.

households. Data are available to the Census block or group level, but I construct the demographic variables at the county level, nearly always using Summary File 3. The main independent variable of interest within the study comes from the Census 2000, and will be discussed in the following section.

The final major source of data is the U.S. Census Bureau's County Business Patterns (CBP). The CBP provide yearly data on the number and size of business establishments within the U.S. Establishment figures are separated at the county level based upon SIC (up through 1997) and NAICS codes (U.S. Census Bureau, 2000a). The only businesses included within the dataset are those with at least one paid employee, so it excludes sole proprietorships and "mom-and-pop" type of stores that are run by one or more individuals that do not receive a paycheck. The CBP is the most reliable source of business establishment data within the U.S. because it is based upon a register of business, and is subsequently updated yearly to be consistent with new business formations.

Nonmonetary Benefits of Self-employment

An important aspect of self-employment returns that the model does not capture because of difficulties in empirical measurement at the county level are the nonmonetary benefits of entrepreneurship. There are many nonmonetary benefits of self-employment, including advantages such as flexible work hours, greater personal autonomy, and higher job satisfaction (McKernan & Salzman, 2008). These benefits are sometimes powerful. Using data from the Survey of Income and Participation (SIPP), Hamilton (2000) finds that workers enter into self-employment even when paid employment offers higher initial earnings coupled with continual higher earnings growth. Over a 10-year timespan, earnings for the self-employed are 35% lower than what could have been earned in paid employment, and very few "superstars" are able to earn more in self-employment than in paid employment. The reasons behind this phenomenon are mainly nonpecuniary in nature, such as "being your own boss" (Hamilton, 2000, p. 628). In coordination with being able to be one's own boss, an individual's preference for independence is also positively related to choosing self-employment (Douglas & Shepherd, 2002). Similarly, the expected utility from employment increases with the associated independence of the job (Douglas & Shepherd, 2002). The above studies show that autonomy, which is difficult to measure, is a major influence in both the decision to become an entrepreneur and the resulting satisfaction of self-employment.

Other researchers have also found that jobs with greater autonomy, even if they are within paid employment, lead to higher compensating returns. Duncan (1976) adds nonpecuniary benefits to the earnings function and finds that having the ability to be flexible in either increasing or decreasing working hours increases the benefits of employment. Duncan and Stafford (1980) extend the idea of greater working freedom to include the ability to choose a flexible work schedule and location of individual work, and again find that greater working freedom leads to positive employment returns.

Greater job satisfaction for the self-employed is another major nonmonetary benefit. Even though self-employed workers tend to work more hours, suffer lower incomes, and are more financially stressed than paid employees, they are still happier with their jobs (Morin, 2009). Nearly 4 out of 10 self-employed workers surveyed are completely satisfied with their jobs, in comparison to under 3 out of 10 who were not self-employed (Morin, 2009). An additional comparison of job satisfaction levels of the self-employed versus employees using the U.S. General Social Survey (Blanchflower, 2004) shows the same results and can be found in Appendix B. Blanchflower and Oswald (1992) also suggest that those who establish and run their own businesses feel happier than those who are working for others. Utilizing a procedural utility approach, Benz and Frey (2008) find that the self-employed value not only the monetary outcomes of self-employment (i.e., income), but also the process of obtaining that income. In other words, the self-employed gain a higher utility from the everyday work they complete when compared to paid employees.

Higher satisfaction for the self-employed remains a strong trend internationally. When comparing self-employment data from 70 countries, Blanchflower (2004) finds much of the same characteristics: even though entrepreneurs are likely to have higher stress levels, be under constant strain, and lose sleep because of their jobs, they are still likely to report that they are very satisfied with their lives. Self-employed workers specifically within Britain also experience significantly higher levels of well-being compared to salaried workers (Blanchflower & Oswald, 1998). Again, it is important to recognize that greater autonomy, greater satisfaction, and flexibility are important considerations in determining returns to employment; however, these characteristics do not lend themselves to measurement at the regional level.

Description of Variables

Dependent Variable: Entrepreneurial Depth (ED)

The method for measuring entrepreneurship affects how the returns to entrepreneurial activity are measured. Entrepreneurship and entrepreneurial activity are difficult to define (O'Farrell, 1986; Gartner, 1990). Past researchers have defined entrepreneurship by using the self-employment rate, business-ownership rate, and firm entry and exit rates (for a review of entrepreneurship measures, see Iversen, Jorgensen & Malchow-Moller, 2008). Within the United States, there are numerous measures that are available at the county level that could be used as a proxy for entrepreneurship (Clayton & Spletzer, 2006). The Quarterly Census of Employment and Wages (QCEW) provides information on employment by size and SIC/NAICS codes at county level. The Business Employment Dynamics, a special tabulation of the QCEW, provides information on firm births and deaths, job creation and destruction, and business expansions and contractions. The economic census administered by the U.S. Census Bureau allows for the creation of the Business Register and the Longitudinal Business Database. The American Community Survey, National Longitudinal Survey, and the Current Population Survey all also provide measurable proxies for entrance into entrepreneurship.

To measure entrepreneurial returns, I use a measure in accordance with Low (2004) and Low, Henderson, and Weiler (2005) coined “entrepreneurial depth”. Entrepreneurial depth (ED) is an important measure of the success of entrepreneurs within a specific geographic context. The entrepreneurial depth variable for county i is as follows:

$$ED_i = \ total\ self - employment\ income_i \div total\ number\ of\ self - employed_i$$

where self-employment is a measure of the number of nonfarm self-proprietors as calculated by the BEA. Total self-employment income is the sum of sole-proprietorship incomes reported on Schedule C of Internal Revenue Service (IRS) form 1040 (Profit or Loss from Business), IRS Form 1065 (U.S. Return of Partnership Income) for partnerships, and the sum of all income received by certain rural and agricultural cooperatives (BEA, 2010). The BEA makes an adjustment for underreporting of incomes by self-employed, which increased nonfarm proprietor incomes by nearly 40% in 2007 (BEA, 2010). I exclude farm self-proprietors’ income from this analysis because of the relatively small contribution to overall self-employment returns (only 5% of the total) and the influence that government subsidies have on these returns.

More entrepreneurial depth within a region leads to higher value added to the local economy. An area may have a high concentration of entrepreneurs, but the type of occupations that they are in varies greatly. Many choose to begin their own business to pursue a dream or partake in activities that they enjoy. For example, a woodworker may decide to sell his wares or a baker may choose to open a small store to sell goods that he or she already enjoys producing. These entrepreneurs do contribute to the regional economy, however marginally. "High-value entrepreneurs," alternatively, have the ability to directly influence to economic growth (Low, 2004, p. 3). These more successful self-employed individuals earn a higher income than those pursuing entrepreneurship as a lifestyle. The high-value entrepreneurs capitalize on their distinct competitive advantage to accumulate wealth, and subsequently have the ability to finance additional new business ventures (Low, Henderson, & Weiler, 2005). Areas with higher concentrations of high-value entrepreneurs will also be the ones with higher entrepreneurial depth.

I use the percent change in entrepreneurial depth between 2000 and 2007 as the dependent variable. Using the percent change allows for the measures to be comparable across counties regardless of population change or inflation. Change in entrepreneurial depth is:

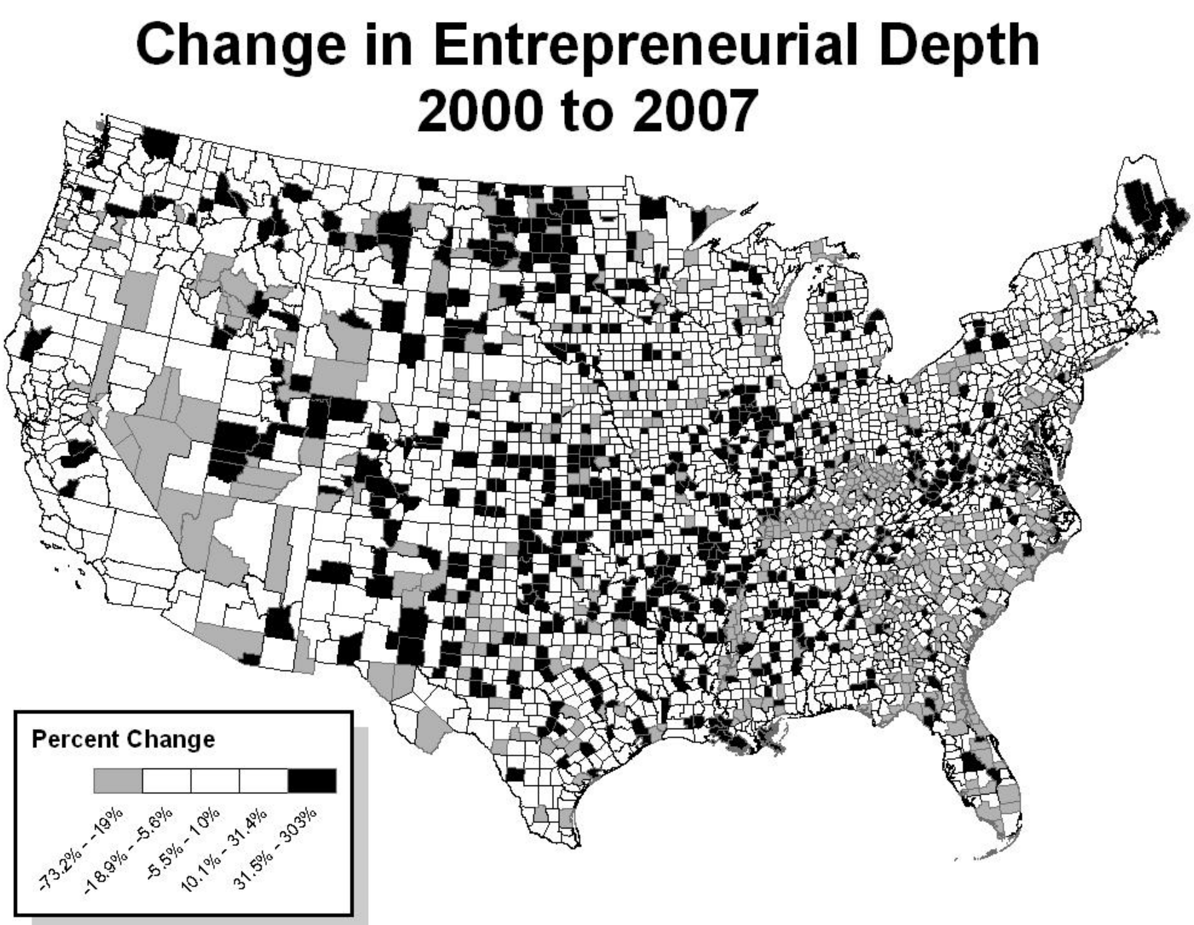
$$\%\,\Delta\, ED_i = (ED_{it+7} - ED_{it}) \div ED_{it}$$

where t is equal to the reference year 2000 and i is an index of each county level unit included within the analysis. Counties with higher percent change in entrepreneurial depth experienced more rapid growth in self-employment returns. Higher entrepreneurial depth contributes to the success of the local economy; hence, counties with high change in ED see positive economic growth as compared to low change counties, *ceteris paribus*.

Regional and Rural Variation of Entrepreneurial Depth

There is significant regional variation in the change entrepreneurial depth. Figure 7 is a map of the top and bottom 20% of counties in terms of change. Black coloration signifies those counties in which self-employment profit increases are in the top quintile. Counties colored gray are those in which self-employment returns were the slowest growing. The highest quintile counties are more geographically dispersed than the lowest quintile counties. There are certain clusters of high positive depth growth counties throughout the Midwest portion of the country, such as in North Dakota, Kansas, Illinois, and Arkansas. Statistically, the Midwest does have the

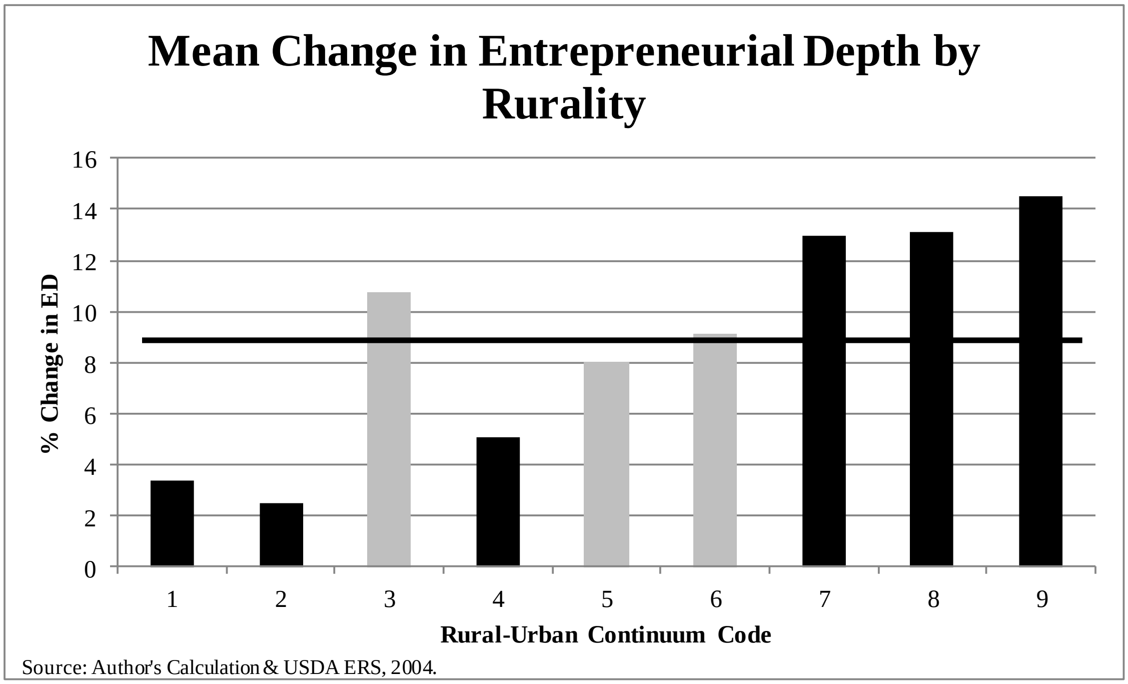
Figure 7: Map of Change in Entrepreneurial Depth



highest mean change among all counties, with an average growth in entrepreneurial depth of around 15%. Conversely, negative returns are more common throughout the Southeast portion of the nation. Georgia, South Carolina, and Florida contain many counties where growth in entrepreneurial depth has struggled. Kentucky, in particular, has seen returns to self-employment decrease in more counties than any other state. Nevada also has a high proportion of poorly performing counties. Interestingly, the South is not the lowest average growth region. The Northeast region is the slowest growing of all regions, with an average gain in income of around 4%. Even though there are a very large number of slow growing counties within the South, there are also some very high growth counties that may be providing a boost. The maximum entrepreneurial depth growth within the South is over 300%. The maximum value of depth growth in that region is 157%. This shows that there may be a macro-level "superstar" effect (Rosen, 1981) within the South region, where a few counties with rapid growth are bringing up the region's overall average.

Change in entrepreneurial depth also varies greatly across levels of rurality. Figure 8 summarizes change in returns to self-employment by each Rural-Urban Continuum code class. The overall mean for all counties is 9.2% (shown by the black trend line). The black columns in the figure show the subgroups where means are significantly different from the overall average. Those in gray are not statistically different from the overall average. It is interesting to find that growth in entrepreneurial income has increased the most rapidly in some of the most rural of areas (codes 7, 8, and 9). Conversely, growth has been slower in the more metro areas (1, 2, and 4) even though they had initially higher levels.

Figure 8: Change in Entrepreneurial Depth by Rurality



Main independent variables of interest: Measures of Inmigration and Outmigration *(M)*

The main interest within this research is to explore whether there is any relationship between domestic population movement and the growth in self-employment profits as measured by entrepreneurial depth. To measure this, the model includes differing measures of internal migration. The main measure of interest is a measure of the dispersion of migration networks at the county level: network entropy. The second measure measures the absolute volume of inmigration and outmigration to or from a county based upon the population at risk of migrating.

Using data from the 2000 U.S. Census of population, I create measures of inmigration to and outmigration from counties within the United States (U.S. Census Bureau, 2000b). The actual migration counts are calculated by the Census Bureau based upon the answer to a two-part "Residence 5 Years Ago" question on the Census long-form. The first part of the question asks

"Did this person live in this house or apartment 5 years ago (on April 1, 1995)?" If the answer to this question is "No, different house in the United States," the individual is asked to provide the city, town, or post office of the residence in 1995. An individual is considered to be a county-to-county migrant if their reported residence in 2000 is located in a different county than their reported residence in 1995. For example, an individual residing in Orange County, New York in 2000 who resided in Chester County, Pennsylvania in 1995 is considered an internal migrant. A person who moved from one household to a different household within the same county, however, is not considered a cross-county migrant. The individual movers are summed to the county level, which leads to the construction of an aggregate county-to-county measure for instances where there was at least one individual move from one county to another. It is important to note that the migration data does not include international migrants to the United States.

Migration Entropy (ε)

The independent variable that characterizes the dispersion of migration networks throughout the U.S. is a network centrality measure termed entropy. It captures the heterogeneity within the migration network. The use of an entropy measure to characterize networks is not a novel idea. Shannon (1948) was the first to develop the calculation of entropy within the field of Information Theory. Shannon's original entropy calculation measures the degree of uncertainty within a system. In many ways, entropy in this reframed migration network typology is still a measure of uncertainty. Tutzauer (2007, p. 253) provides a detailed description of the relation between uncertainty and information within a system:

> "When choices are highly constrained, as when it is very likely that certain symbols will be sent (or when a flow is highly likely to end at just a few nodes),

> then information (and centrality) is low. When there is much freedom of choice, as when all symbols are equally likely to be sent (or a flow is equally likely to stop anywhere), then information (and centrality) is high."

Within the context of migration networks, "symbols" can be related to migrants and "nodes" to counties. When there is a high likelihood that a migrant will "end," or move to, just a few counties, there is low centrality and information. This explanation tells us that counties that have predictable migration flows to or from them possess lower centrality, and thus are less likely to be hubs for diverse information flows. Conversely, when there is a relatively equal likelihood that migrants will move to/from a county, centrality and information content is high. Counties with unpredictable migration flows to or from them possess higher centrality, which positions them in a more advantageous place for sharing heterogeneous information within the network.

For a more general understanding of how the entropy measure relates to migration networks, it is helpful to build an example of how movement may occur. The node in the context of this analysis is the county, and the specific object being passed from one node to another is the individual migrant. If an object begins at a highly central node, it will be difficult to determine where it will finish within the network, indicating the presence of high entropy. In the context of migration networks, there is a high degree of uncertainty as to where a migrant from a high outmigration entropy county will migrate to. In other words, migrants originating in a county with high outmigration entropy have a wider variety of locations to where they move; hence, the network originating from the origin county will be more dispersed and heterogeneous. A more heterogeneous migration pattern increases the breadth of potential information flow back to the home county. The case is similar for a county with high inmigration entropy. There is a high degree of uncertainty as to where migrants are coming from, indicating that inmigration is taking place from a diverse spread of counties. Migrants going into the high entropy counties come from varied locations, carrying with them diverse localized knowledge. With this variation comes a wider breadth of information flows. Counties that have high values of in- or out-migration entropy can be termed “hubs” of migration, or those with the most diverse migration streams and varied information flows.

Contrastingly, if an object begins at a non-central node, it is relatively easy to assess where it will end within the network. A relatively high probability of the object ending in a specific node or set of nodes indicates the presence of low entropy. Linking this to migration networks, counties with a low degree of uncertainty as to where most of the outmigrants go will also have low outmigration entropy. Outmigrants from low outmigration entropy counties are more likely to end up in the same destination county, which leads to a highly homogeneous pattern of migration and weaker centrality. The lower centrality limits the range of information flows back to the origin counties. The case is the same for counties with low inmigration entropy: migrants coming in are from a homogeneous set of counties, which restricts where ideas are flowing from and subsequently reduces variety and diversity of information content.

Researchers in the past have used entropy measures to help define diverse economic activities, but no one has applied this measure strictly to migration networks. Semple and Golledge (1970) use entropy to test the uniformity of urban settlement patterns in Canada. Since the 1970s, regional economists have used an industrial entropy index to determine the concentration of employment or income in a particular sector (Siegel, Johnson, & Alwang, 1995). More recently, researchers have used Shannon’s entropy calculation as a measure of product variety in a market (Straathof, 2007). Eagle, Macy, and Claxton (2010) create entropy values to measure the diversity of communication networks within the UK. Goetz et al. (2010) were the first to apply the network entropy measure to detailed U.S. county level data. The authors generate entropy measures for both in- and out-commuting networks of U.S. to explore the relationship between this measure of centrality and regional income growth. The use of entropy centrality within the network framework is a relatively new application, and by exploring how these measures may affect self-employment profits adds to the growing literature in the area.

For the calculation of inmigration and outmigration entropy, I use the same form as Goetz et al. (2010), with the exception of substituting migration networks for commuting networks. Suppose that m_{ij} is the actual number of individuals who lived in county i in 1995 and migrated to county j by 2000. Inmigration entropy *(inentropy)*, or the index of diversity of information content of a county's inmigration network, is calculated as follows:

$$p_{ij} = m_{ij}/\sum m_{ij}$$

$$\varepsilon_j^{in} = -(\sum_i p_{ij} \log_2 p_{ij}) / \log_2 N$$

where p_{ij} is the weighted probability that county j receives a migrant from county i during the time period in question, the sum of p_{ij} includes all counties that send migrants, and N is the total number of nodes within the network (equal to 3,047, the number of U.S. counties included within the analysis). It is important to note that the calculation relies upon there being at least one inmigrant to a county, which is not an issue for data within the analysis because all counties have at least one inmigrant. Dividing by the log of the number of nodes within the network normalizes the entropy measure to make it comparable across networks of different sizes. Outmigration entropy *(outentropy)* is calculated using the same formulas, except for the reciprocation of the subscripts[8]. Even though there are separate measures for inmigration entropy and outmigration entropy, they both capture the same essence. Inmigration entropy is a measure of the dispersion

[8] For outmigration entropy, all of the same conditions hold as do for inmigration entropy. Let m_{ji} equal the actual number of people who lived in county j in 1995 and migrated to county i by 2000. The weighted probability that a county j sends people to county i is:

$$p_{ji} = m_{ji}/\sum m_{ji}$$

The subsequent outmigration entropy measure for county j is:

$$\varepsilon_j^{out} = -(\sum_i p_{ji} \log_2 p_{ji}) / \log_2 N$$

of the network of migrants that has migrated into the county. Outmigration entropy, conversely, measures the dispersion of the migrant network that has left the county over the previous five year timespan, effectively indexing the extent of diversity of potential information flows within the network.

Instead of examining only the absolute number of inmigrants or outmigrants that a county has had over the past years, the entropy variable takes into account the variety of places that those migrants are going to and how many are going to each location. Entropy may be relatively higher for a county that sent few migrants to a very diverse set of counties than for a county that sent a larger number of migrants to a more homogeneous set of counties. Data from Billings County, North Dakota and Storey County, Nevada show how this can happen in reality. Billings County lost 191 people to outmigration between 1995 and 2000, while Storey County lost a total of 537. In conventional measures of centrality such as outdegrees, Storey County would be expected to possess the larger network and higher centrality because of the larger number of absolute outward linkages. Using the measure of entropy, however, there is minimal difference between the centrality of these two counties because the diversity of destination locations is taken into account. The outmigration entropy value for Storey County is .2476, while the same value for Billings County is .2427. The difference between the entropy measures is minimal in comparison to the actual number of outmigrants because of the diversity of destination counties. The small number of outmigrants from Billings County are dispersed throughout 17 counties, while the much larger number of outmigrants from Storey County are dispersed throughout only 19 counties. The outmigrants from Storey County are also more heavily concentrated among the most popular destination counties. For example, 75% of migrants from Storey County relocate to the top 4 most popular counties. Only 70% of Billings County's outmigrants move to the 4 highest density destinations. The calculations behind these data and the construction of the entropy measures for the two counties are provided within Appendix D.

Regional and Rural Variation Among Entropy Measures

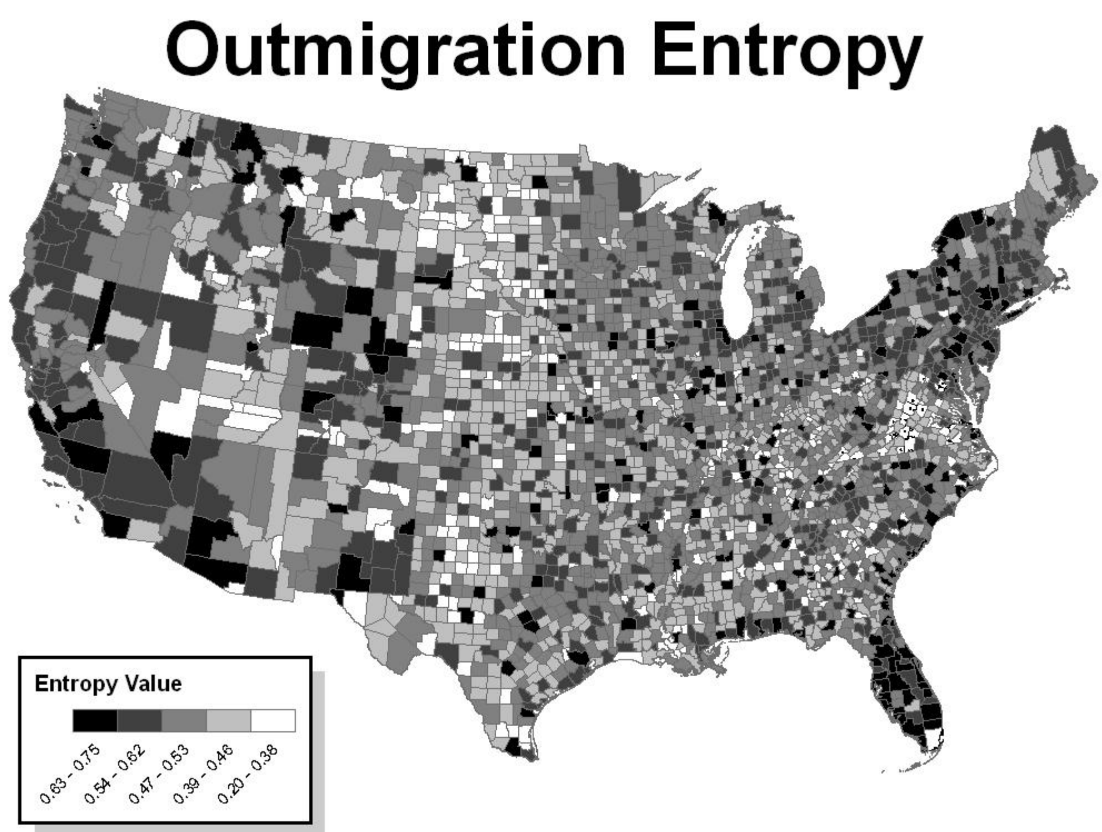
There is significant regional variation in the levels of inmigration and outmigration entropy. Summary statistics for the entropy measures by region of the U.S. provided in Table 6 offer some insight into the existing differences across space. Inmigration entropy is fairly similar across all regions, with the highest mean found within the Northeast and the most within-region variation within the South (both the minimum and maximum scores are found within this region). Outmigration entropy scores, conversely, are more differentiated across regions. The Northeast and the West have the highest average score, with the South again characterized by the most variance. It is interesting to note that counties within the two regions of the United States where outmigration has been an enduring trend, the Northeast and Midwest, tend to have much different outmigration entropy scores. Most counties within the Northeast have higher levels of entropy, while those in the Midwest tend to have lower entropies. This suggests that migrants leaving Northeast counties generate more dispersed networks by which more information can flow when compared to Midwest counties.

Table 6: Entropy Summarized by Region

Entropy Measures Summarized by Region of the U.S.					
	Number of Counties	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
Inmigration Entropy					
Northeast	217	0.5014	0.0670	0.2854	0.6969
Midwest	1055	0.4825	0.0703	0.2901	0.7057
South	1362	0.4934	0.0852	0.2199	0.7783
West	413	0.5009	0.0793	0.2626	0.7536
Outmigration Entropy					
Northeast	217	0.5398	0.0627	0.3573	0.7095
Midwest	1055	0.4792	0.0719	0.2425	0.6837
South	1362	0.4798	0.0838	0.2022	0.7523
West	413	0.4934	0.0886	0.2403	0.7354
Source: Author's Calculations					

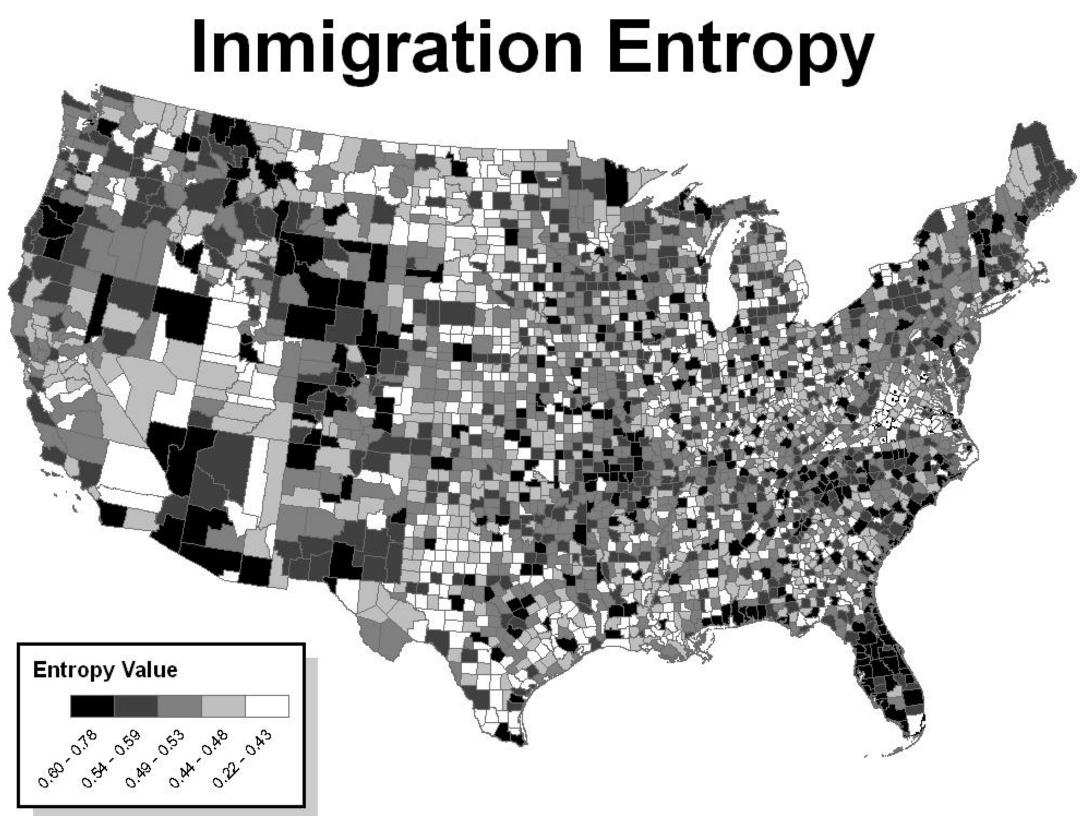
Figure 9 and Figure 10 are mappings of the outmigration and inmigration entropy measures at the county level by quintile. It affords a more detailed insight into regional variation. The darkest of counties demarcate the highest entropy scores, with the lightest counties showing the lowest entropy scores. Examining the outmigration entropy map, it is clear that there are certain regions where there are clusters of high and low value counties. Florida, for example, has a heavy concentration of high entropy counties. Migrants from these counties move to many other different counties, creating a largely dispersed migration network. Other areas that have high outmigration entropy counties include Chicago and its suburbs, the New York City metro area, and the San Francisco/Silicon Valley area.

Figure 9: Map of Outmigration Entropy



Low outmigration entropy scores appear to be highly concentrated within the Midwest (as expected from the discussion above), especially the counties of North and South Dakota. There are also groups of counties within the Appalachian Region, Utah, and Louisiana that experience lower levels of entropy. Migrants from these areas tend to have migration networks that are restricted to a small set of places.

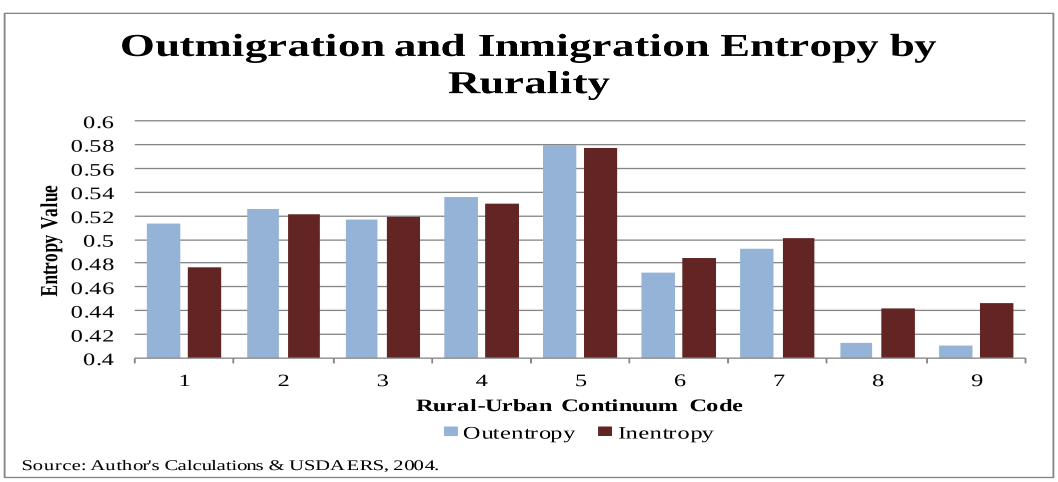
Figure 10: Map of Inmigration Entropy



It is more difficult to make regional generalizations about inmigration entropy. There are fewer distinct clusters throughout the U.S. The most distinct clustering lies within Florida, which also has high outmigration entropy. Other areas that have high inmigration entropy include the Eastern Tennessee and Western North Carolina region, Southwest Arizona, Colorado, and areas around the Ozark Mountains. Low inmigration entropy, like outmigration entropy, is centered throughout the Midwest in the Dakotas and parts of Texas. Utah and parts of Nevada are also clusters of low entropy counties. It is interesting to see that areas around Silicon Valley, which have high outmigration entropy, tend to have lower inmigration entropy.

There is also significant variation among the entropy measures when categorized by rurality. Figure 11 shows entropy levels broken down by the ERS's Rural-Urban continuum. When looking at overall entropy, or both in and outmigration entropies, metro areas tend to have higher levels of entropy than nonmetro areas, with the exception of inmigration entropy values of the most urban of places, where the average value is surprisingly low. Counties with an urban population of 20,000 or more that are not adjacent to a metro area have the highest levels of both in- and outmigration entropy. The more metro areas also have similar levels of entropy when examining both inflows and outflows, with the exception again being the most urban of places, where outmigration entropy is significantly higher than inmigration entropy. The most rural areas also have highly differentiated levels of entropy. Counties coded with either an 8 or 9 by the ERS attract migrants from a wider variety of places and have more highly dispersed inmigrant networks than outmigrant networks. Migrants from these counties do not go to a wide variety of places, which suggests that they are limited by a certain common characteristic or set of characteristics.

Figure 11: Entropy by Rurality



Construction of Average Entropy (avgent)

Within the models, I average inmigration and outmigration entropies and use this as the independent variable that proxies diversity of migration networks *(avgent)*. The average is a simple calculation:

$$Average\ Entropy_j = \frac{\left(\varepsilon_j^{in} + \varepsilon_j^{out}\right)}{2}$$

where ε_j^{in} is inmigration entropy, ε_j^{out} is outmigration entropy, and j is a county index. The main reason to use average entropy instead of separate measures for outmigration and inmigration is that it facilitates the construction of a single model. Because inmigration and outmigration entropy are highly correlated with one another, it becomes more difficult to assess each measure's contribution to change in entrepreneurial depth if both are included[9]. Instead, average entropy serves as a proxy for counties with both high or low in- and outmigration entropy. Average entropy also picks up on a county that may have a high value of inmigration entropy combined with a low value of outmigration entropy or vice versa. A county in this position still benefits from diverse knowledge flows from either its outmigrants in differentiated locations or its inmigrants from a wider breadth of counties. One drawback of using average entropy is that the effects of either high or low entropy outlier counties are mediated, potentially influencing results. Results of regressions using average entropy (discussed further in Chapter 5) are very similar to those obtained by running two separate regressions for in- and outmigration entropy[10], thus it is an appropriate alternative that simplifies and clarifies the model. Table 7 provides information on counties with the highest and lowest values of average entropy. The highest values are concentrated within metro areas, while the lowest values are centered in the most rural of areas.

[9] The correlation between outmigration entropy and inmigration entropy is .8191.

[10] For comparison with results obtained by using average entropy values, results of separate regressions for outmigration and inmigration entropy can be found within Appendix C.

Table 7: Top and Bottom 10 Counties, Ranked by Average Entropy

	Lowest Values of Average Entropy	
Value	*County, State*	*R-U Code*
0.299	Banner, NE	9
0.290	Wirt, WV	3
0.287	McPherson, NE	9
0.284	Camas, ID	9
0.283	Glassock, TX	8
0.280	Juab, UT	2
0.279	Echols, GA	3
0.275	Billings, ND	9
0.269	Issaquena, MS	9
0.211	Kenedy, TX	9
	Highest Values of Average Entropy	
Value	*County, State*	*R-U Code*
0.765	Onslow, NC	3
0.747	Cumberland, NC	2
0.744	El Paso, CO	2
0.726	Brevard, FL	2
0.725	Maricopa, AZ	1
0.725	Lee, FL	2
0.719	Bell, TX	2
0.718	Shelby, TN	1
0.718	Comanche, OK	3
0.718	Beaufort, SC	5

Source: Author's calculation, U.S. Census Bureau, USDA ERS

Migration per Capita (inmigPC, outmigPC)

The second measures indicate migration intensity to and from a county in relation to the population base. To make migration streams comparable across counties, I divide the absolute number of inmigrants or outmigrants to or from a county by the population at risk of migrating to form rates. This takes into account different population levels among the counties, and places migration on the per capita basis. It is necessary to normalize the figures in this way to avoid misleading results. For example, there were a total of 96 outmigrants from Arthur County, Nebraska, comprising over 22% of the county's population. If the same number of migrants were to leave a county with a much larger population, such as Lancaster County, Pennsylvania (those 96 migrants leaving would account for only .02% of that county's population loss) the impact on the origin county would be much less.

The population at risk of outmigration from a county is simply that county's population in 1995, before the "Residence 5 Years Ago" question picks up any movers. Counties that have a larger population are more likely to have a larger number of outmigrants. By dividing the number of actual outmigrants from one county by the total population of that county, it is possible to find a per capita rate that is comparable for counties of different sizes[11].

Inmigration per capita, on the other hand, is more difficult to define. The true population at risk of migrating into a particular county is not simply the population of the county at the beginning of the time period. The actual population at risk of inmigration to a particular county is the total U.S. population in 1995 that does *not* reside within that county:

$$\textit{Population at Risk of Inmigration to County } i = \text{U.S. pop}_{1995} - \text{population}_{i1995}$$

The pool of potential inmigrants to Los Angeles County, CA (population of over 9 million) is much smaller than the potential pool of inmigrants to Kenedy County, TX (population of 420). Even though the potential pool is larger for Kenedy County, it does not necessarily mean that there will be more inmigrants there than in Los Angeles County. Dividing the total number of inmigrants by the population at risk for inmigration creates a demographically accurate method for comparing counties.

[11] The final calculation of the outmigration rate for each county is thus:
Outmigration rate = Absolute number of outmigrants$_{1995\text{-}2000}$ / Population at Risk of Outmigration

It is also possible to calculate the inmigration rate in a different fashion. Instead of using the population of all other counties as the population at risk of migrating, past researchers have used simply the population of the county as a base (Shryock & Siegel, 1967; Hunter, 1998). Both measures have their own merits. The first described is a more accurate demographic representation of inmigration because of the use of the appropriate population base. The second calculation is practical in the fact that it lends insight as to the attractiveness of a county to migrants. If there is a high level of inmigration per capita, it is clear that there is one or a combination of characteristics within the county that make it a favorable migration destination. I choose to use inmigration per capita based upon the county's 1995 population because of an issue with multicollinearity when using the total U.S. population as a base[12].

Inmigration and outmigration rates are very similar to another type of network centrality measure noted earlier: Freeman degrees. In the context of migration networks, indegrees would provide a figure equivalent to the absolute number of migrants entering into the county (the node in question) during the 1995-2000 time period. Outdegrees would provide a figure as to the number of migrants that left the county during the five year time period. Because the in- and outdegree measures would be analogous to the absolute migration numbers provided by the Census Bureau data, I choose to use the per capita based figures to take into account different population levels of counties.

Data Limitations

There are certain limitations to the U.S. Census data used to calculate the migration measures. First, the measure captures all migrants that are over 5 years of age. For this research,

[12] Self-employment in 2000, another control variable within the analysis, is highly correlated with migrants per capita with total U.S. population as the base (.861); therefore, using the 1995 county population as the base reduces inaccuracies in related variable standard errors.

this is problematic because those who are not of working age are counted as migrants even though they will likely not influence self-employment income growth directly. They may indirectly influence self-employment if the gain or loss of a younger population affects what new businesses enter or exit the market, but it is difficult to assess the extent of this influence. It is also less likely that this young population will be a main channel by which novel knowledge flows between counties. Data on county-to-county migration by age, which would show the number of young migrants not affecting the workforce, are not available as a special tabulation from the 2000 Census. An alternative source, the CPS, does contain data on movers for young ages for the time period between 1990 and 1995. Using this source, it is possible to see the potential impact of including migration of non-working age individuals (assuming that migratory patterns are fairly similar for the 1990-95 and 1995-2000 periods). The CPS data reveal that approximately 17% of migrants over the age of 5 fall within the 5-14 year old range. These migrants are likely not working, and thus will not directly affect growth in entrepreneurial depth. How this affects the analysis directly is unclear because there is likely to be spatial variation among counties, along with the fact that the CPS data measurement is of gross migration – which does not distinguish between inmigrants and outmigrants. It is, however, important to recognize that this may influence results. A full table showing population movement by age between 1990 and 1995 can be seen in Appendix E.

Secondly, the Census data do not incorporate multiple moves that may have taken place during the five preceding years. It only takes into account the initial and final locations of the population. The mover from Chester to Orange County in the example above may have only moved that one time, but it is also possible that the migrant made more than one move at various points throughout the five years. For example, he or she may have moved from Chester to Centre County, Pennsylvania, to Miami-Dade County, Florida, to Mills County, Texas, to Lake County, Colorado, and finally to Orange County. The multiplicity of moves complicates the understanding of how movement may affect self-employment income growth. If there were multiple moves over the time period, the network created by the migrant would be larger than if only one move occurred. This may create information and knowledge flows that are not picked up by the static nature of the Census question. In this case, the importance of the expanded migration network for entrepreneurial depth would be understated.

Finally, the measure does not take into account the temporality of the move. The move may have taken place at the very beginning of the time period, such as on June 1st, 1995. Conversely, it may also have taken place at the end of the time period, such as on December 20th, 1999. It is impossible to glean from the data how long a migrant has been in the beginning or ending location. The length of tenure in each location may have affected the creation of important personal and business ties that facilitate entrepreneurial success. Length of residence in origin and destination may also affect the extent of the knowledge that can be passed to those in each location. Given the data's drawbacks, it is still the best measure to represent domestic migration within the U.S. for the given time period.

Additional Explanatory Variables

Initial Level of Entrepreneurial Depth (ED_i)

It is necessary to include initial entrepreneurial depth in the base year as a control variable because of the use of a change model. Assuming convergence (see Durlauf, Johnson, & Temple, 2005; Higgins, Levy, & Young, 2006), counties with higher levels of initial entrepreneurial depth will grow more slowly than counties with lower initial returns to self-employment. Initial entrepreneurial depth should have a negative relationship with change over the time period.

County and Economic Indicators (C)

The number of entrepreneurs within the economy may influence entrepreneurial depth. Another measure termed by Low, (2004) which measures the density of entrepreneurs within an economy, is entrepreneurial breadth *(breadth).* This measure is calculated by taking the number of self-employed within the economy and dividing it by the total employment of the county[13]. High levels of entrepreneurial breadth signify a high concentration of self-employed. Low, Henderson, and Weiler (2005) find that high entrepreneurial breadth counties are located in smaller, rural counties where entrepreneurial depth is low. Conversely, low entrepreneurial breadth counties are concentrated in urban areas where entrepreneurs are likely to target higher-value self-employment. Even though a high density of entrepreneurs may create knowledge spillover effects in some regions, there may also be negative effects on profits because of the competition that it creates. If an entrepreneur is reaping profits, others will see this market signal and follow in his or her footsteps, raising entrepreneurial breadth levels (Kirzner, 1973). Subsequently, the new group of self-employed begin to compete away the opportunity for profits by way of increased competition (Glancey & McQuaid, 2000). In agreement with prior research and theory, entrepreneurial breadth and depth should have a negative association.

Risk or uncertainty in self-employment income can also be influential in determining the growth of entrepreneurial depth. Entrepreneurship is inherently risky because of the uncertainty involved in undertaking a new endeavor (Knight, 1921; Parker, 1996; Kihlstrom & Laffont, 1979). The decision to become an entrepreneur may be an inherently risky behavior, but it may lead to increased profits. The model includes a proxy for risk *(risk)*, the coefficient of variation of self-employment income between the years 1995-2004, which has previously been shown to affect entrepreneurial depth (Goetz & Shrestha, 2009). The measure for county i is calculated as:

[13] Self-employed here are again nonfarm self-proprietors as counted by the BEA.

$$risk_i = \frac{\sigma_{ED_{i95-04}}}{\mu_{ED_{i95-04}}}$$

where σ is the standard deviation, μ is the mean, and $ED_{95\text{-}04}$ is entrepreneurial depth between the years of 1995 and 2004. Higher levels of uncertainty symbolize a rapidly changing, dynamic economy. Entrepreneurs can take advantage of this uncertain climate by partaking in risky endeavors that can lead to high profits. As with other aspects of economic theory, higher risk associated with an investment leads to the potential for larger rewards. Risk is thus hypothesized to be positively associated with entrepreneurial depth.

Access to capital is another important consideration that determines success of entrepreneurs. Financial capital is an important factor in entrepreneurial development (Dunn & Holtz-Eakin, 2000; Holtz-Eakin, Joulfaian, & Rosen, 1994), but it also has a positive impact on enduring entrepreneurial success (Cooper, Gimeno-Gascon, & Woo, 1994; Bates, 1990). Regions with entrepreneurs that have elevated access to financial capital will likely have more successful high-value entrepreneurial endeavors. Local high-value entrepreneurs create a self-perpetuating cycle of greater access to capital through higher profits, which can be utilized to invest in future ventures (Low, Henderson, & Weiler, 2005). As a representation of a region's access to financial capital, I use bank deposits per capita in 2001 *(deposits)*[14]. Data on total bank deposits within a county are available through the Federal Deposit Insurance Corporation (FDIC, 2004). Banking deposits per capita are total bank account deposits in 2001 divided by the U.S. Census Bureau's population estimate for the same year. This measure should be positively associated with entrepreneurial depth.

The number of educational establishments and business support establishments should impact entrepreneurial depth (Goetz & Shrestha, 2009). Examples of business support establishments include big box office supply stores that offer copy and editing services, temp

[14] Data on total bank deposits were not available for the reference year, so the next closest year available was used.

agencies that provide temporary workers during business intensive seasons, and computer stores that provide hardware and software assistance. Business support establishments provide services to entrepreneurs that they would originally have had to complete on their own. They save time and allow the entrepreneur to focus on their core competitive advantage. Following Goetz and Shrestha (2009), I construct the number of business support establishments within the county *(bsestab)* by summing the total number of establishments in the following NAICS categories: couriers and messenger services (492), office supply and stationary stores (453210), computer and software stores (44312), business support services (5614), temporary help services (561320), and child day-care providers (624410). Educational establishments *(edestab)* are facilities that cater to training entrepreneurs. They include junior colleges (NAICS 611210), business schools and computer management training (6114) and technical and trade schools (6115) (Goetz & Shrestha, 2009). These establishments help to train the self-employed in a wide variety of disciplines, including basic business management skills, technological advances, and trades necessary for starting a novel business. For each measure, I divide by the county population in 2000 to show relative accessibility to the services offered. Access to business support and educational establishments should increase returns to self-employment and help to propel positive change in entrepreneurial depth.

The presence of a large, creative class of individuals within a county is another important aspect of entrepreneurial depth. Creativity fosters entrepreneurial activity within a region and is an important source of innovation, especially in high-technology sectors (Lee, Florida, & Acs, 2004). The high-tech industrial sector creates high-value entrepreneurs that drive economic growth. In past research, a high concentration of people employed in arts and entertainment industries has led to higher returns to self-employment (Low, Henderson, & Weiler, 2005), confirming that high creativity fosters high-value entrepreneurs. Florida’s (2002a) Bohemian Index is an indicator for the creative class of a region. I create the index based upon creative class county codes generated by the USDA ERS (McGranahan & Wojan, 2007). The bohemian index *(bohemian)* is a "location quotient that measures the percentage of bohemians in a region compared to the national population of bohemians divided by the percent of population in a region compared to the total national population" (Florida, 2002a, p 59). It is constructed as follows:

$$bohemian_i = \frac{\frac{Emp_{Arts_i}}{Emp_{Arts_{U.S.}}}}{\frac{Emp_{Tot_i}}{Emp_{Tot_{U.S.}}}}$$

where employment in the arts includes the number of artists and design workers (Standard Occupation Code (SOC) 27-1000) and entertainers and performers, sports, and related workers (SOC 27-2000). I use total employment as the denominator of the equation because it accurately represents the county's workforce which the bohemians are part of. The bohemian index is theoretically positively associated with entrepreneurial depth.

Other county level characteristics that may influence entrepreneurial depth are the population density and the percent of the workforce in construction employment. Population density of the county should be positively associated with higher returns to self-employment and entrepreneurial capital, especially in certain industry sectors such as high tech and information and communication technology (Audretsch & Keilbach, 2004). These industries are also those that are highly likely to spawn high-value entrepreneurs that drive economic growth within the region. Having a dense concentration of people also creates a deeper product market. Even if two entrepreneurs have the exact same levels of physical productivity, the one with a larger number of potential customers close at hand will be more successful because of reduced transportation and marketing costs (Venables, 2006). The population density *(density)* of a county is simply the population divided by the land area. Construction is one industry sector that lends itself well to self-employment (Hundley, 2000). Many times, home builders are self-employed, and they subsequently subcontract out parts of the housing construction process to other self-employed specialists in plumbing, electricity, and the like (Goetz & Shrestha, 2009). Construction may be amenable to the generation of self-employment jobs, but it does not necessarily lend itself to the type of high-value entrepreneurship that is the driver of economic growth. It is less likely that individuals involved in construction trades will begin novel entrepreneurial endeavors that create high value entrepreneurship due to the unique skill set that is associated with construction. I use BEA data on construction and total employment to create the proportion of the county workforce that is employed in the construction sector *(cons)*[15]. I hypothesize that higher levels of construction employment will be negatively associated with entrepreneurial depth.

Finally, I include two dummy variables to test whether there are differences among nonmetro/metro counties (*nonmetro*) and counties that either lost or gained population to migration during 1995-2000 (*netmigdum*). The USDA ERS' 2003 county classifications determine the metro/nonmetro status of a county. Counties that lost population to migration are those where net migration is negative (Net Migration = Inmigration – Outmigration). I include this variable to see if growth in self-employment incomes is lower in counties that are losing population to migration. I control for this in addition to in- and outmigrants per capita because a county with a high proportion of inmigrants (outmigrants) may still have net negative (positive) migration if there is also a counteracting high proportion of the population leaving (inmigrating). I expect counties with negative net migration to experience lower growth in entrepreneurial profits due to assumed population loss. By including these variables, it is possible to see whether

[15] Due to confidentiality issues, the BEA does not report construction employment where the number of people employed in the sector is very low. Construction employment in these counties is not assumed to constitute a significant proportion of employment within the county, and thus is assumed to be zero percent of the total workforce.

entrepreneurial depth grew faster, slower, or at the same rate when controlling for other county level characteristics.

Human Capital and Experience indicators (H)

Enhanced human capital and greater levels of experience theoretically influence returns to entrepreneurship positively. The first experience variable is the median age *(medage)* of the population[16]. This variable serves as a proxy for experience in both paid and self-employment. As an individual becomes more experienced in business functions by partaking in either wage and salary employment or self-employment, he or she is expected to accumulate skills that allow for taking advantage of potentially profitable business ventures (Duchesneau & Gartner, 1990). Once an individual enters self-employment, it is also expected that more experienced individuals will be more effective in running the business, effectively raising productivity and increasing income (Evans & Leighton, 1989). Model specifications include a quadratic term *(medagesq)* to account for the self-employment returns of older individuals who may have supplemental social security income. Even though elder entrepreneurs may still have the experience necessary to effectively run a more profitable business, high amounts of income from entrepreneurial activity may not be needed because of government-provided social security payments (Fuchs, 1982). It may also be the case that to receive retirement benefits, individuals must at least partially retire, which would reduce the amount of income gained from self-employment activities. Hamilton (2000) finds that experience, measured by potential labor market experience, has an inverse U-shaped relationship to self-employment earnings. Hence, the expected sign of the linear median age variable is positive, with the opposite expected for the quadratic median age term.

[16] The U.S. Census Bureau generates the median age for each county population in file P013001 of the Census 2000 Summary File 1.

Educational attainment theoretically influences entrepreneurial depth, as discussed in Chapter 2 above. I construct two variables to measure different levels of human capital accumulation. The first is the percent of the adult population (age 25 and over) that has obtained a college degree *(college).* The second is the percent of the adult population that has not earned a high school diploma or equivalent *(hsdeg)*[17]. Having a college education is theoretically positively related to higher change in entrepreneurial depth, and not having completed high school should be negatively associated with returns to self-employment.

Demographic Indicators (D)

Based upon past research and theory, the analysis includes a number of demographic control variables that influence entrepreneurial success. One potential factor that influences returns to self-employment is racial composition of the population. Minorities experience both lower rates of and returns to self-employment than whites due to consumer discrimination (Borjas & Bronars, 1989; Meyer, 1990); therefore, a county with a high proportion of minorities would expect to experience lower entrepreneurial depth growth. I use the ethnic fragmentation index introduced by Alesina, Baqir, and Easterly (1999) to capture the degree of racial heterogeneity of the population *(ethnic)*. The ethnic fragmentation index is as follows:

$$\textit{Ethnic Fragmentation Index} = 1 - \sum_i (Race_i)^2$$

[17] Educational attainment data is available through the Census 2000 Summary File 3. I construct the college education variable by summing the total number all of those who have achieved higher than a bachelor's degree and dividing by the total population. The use of specific census files is as follows: ((P037015:P037018+P037032:P037035)/P037001). The high school education variable is constructed by summing the number of all of those below obtaining a high school diploma and dividing by the total population. The use of specific census files is as follows: ((P037003:P037010+P037020:P037027)/P037001).

"where $Race_i$ denotes the share of the population self-identified as race i ∈ I = {White, Black, Asian and Pacific Islander, American Indian, and Other}" (Alesina, Baqir, & Easterly, 1999, p. 1255). The index measures the probability that two people who are randomly selected from a county's population will be of the same racial category[18]. Higher levels of ethnic fragmentation will theoretically lead to lower entrepreneurial depth.

The gender composition of the workforce also has the potential to influence returns to self-employment. Within the U.S., women's participation in non-professional self-employment such as childcare harms earnings (Budig, 2006). Using data from Germany, Georgellis and Wall (2005) find that women who entered self-employment experienced significantly lower incomes for the first year. Even though entrepreneurship may be a solution for women to escape from the "glass ceiling" in paid employment (see Buttner & Moore, 1997 for discussion), switching to self-employment does not always mean that incomes for women are the same as for men in self-employment. Women in Italy are also less likely to succeed in entrepreneurial endeavors (Rosti & Chelli, 2005), which indicates that the income produced during self-employment is low. Based on this research, a higher proportion of women in the county's labor force *(female)* should lead to lower entrepreneurial depth[19].

The presence of a high concentration of foreign-born people in the economy may have an impact on the county-level earnings profiles of the self-employed. While there is some evidence in other fields, such as sociology, that immigrants are successful within entrepreneurial business ventures, economists' findings show little support for this statement (Portes & Zhou, 1996). Borjas (1990, p. 163-164) states: "the presumption that many immigrant entrepreneurs begin with small shops, and with their ability and hard work accumulate substantial wealth is a myth." By

[18] The U.S. Census Bureau provides the number of people in each county that self-identify themselves as a certain race. Each race is provided in a separate file, P007001-P00708, of the Census 2000 Summary File 1.

[19] The percent of the labor force that is female is calculated using U.S. Census Bureau's Census 2000 Summary File 3. The total number of females in the labor force (P043010) is divided by the total labor force (P043003+P043010) to determine the percent female.

the late 1990s, immigrants were actually less likely to be self-employed than their native counterparts, a reversal of historical trends dating from the 1960s (Camarota, 2000). This signals that profitability and success of self-employment is substantially low enough as to not attract additional foreign-born entrepreneurs. To control for the immigrant population, I use the percentage of the total county population that is foreign born *(foreign)*[20]. The expected sign of the variable is negative due to past findings that immigrant entrepreneurs' success is limited.

Descriptive statistics on each of the variables within the model can be found in Table 8. The average change in entrepreneurial depth over the time period is over 9%, with Clark County, Idaho experiencing the most severe decline in self-employment income (73%) and Osceola County, Michigan experiencing the largest increase in income over the time period (303%).

[20] The U.S. Census provides the total foreign born population of the county in Summary File 3.

Table 8: Descriptive Statistics

Descriptive Statistics				
Variable	**Mean**	**SD**	**Min.**	**Max.**
Percent Change in Entrepreneurial Depth (EDchg), 2000-07[a c]	9.159	34.044	-73.223	303.039
Average Entropy (avgent), 1995-2000[b c]	0.488	0.076	0.211	0.765
Inmigration Entropy (inentropy)*, 1995-2000[b c]	0.491	0.079	0.220	0.778
Outmigration Entropy (outentropy)*, 1995-2000[b c]	0.486	0.081	0.202	0.752
Number of Outmigrants (outmigPC), 1995-2000 (per 1995 County Population)[b]	0.188	0.059	0.085	1.034
Number of Inmigrants (inmigPC), 1995-2000 (per 1995 County Population)[b]	0.196	0.077	0.047	0.803
Initial Entrepreneurial Depth (ED2000), 2000 ($1,000)[a]	19.174	7.973	3.929	144.344
Entrepreneurial Breadth (breadth), 2000[a c]	19.330	6.638	1.590	56.360
Coefficient of Variation of Entrepreneurial Depth (risk), 1996-2005[a c]	21.294	9.628	4.342	67.600
Deposits per Capita (deposits), 2001 ($1,000)[b d]	11.864	7.676	0.000	201.946
Educational Establishments (edestab), 2000 (per 10,000 population)[f]	0.245	0.371	0.000	4.766
Business Support Establishments (bsestab), 2000 (per 10,000 population)[f]	4.271	2.473	0.000	26.662
Population Density (density), 2000 (persons per square mile)[b]	218.780	1661.551	0.271	66940.080
Bohemian Index (bohemian), 2000[e c]	0.596	0.389	0.000	5.741
Construction Employment (cons), 2000 (% of total employment)[a]	5.661	3.043	0.000	30.053
Nonmetro Dummy Variable (nonmetro), 2003, (1 if nonmetro)[g]	0.658	0.475	0.000	1.000
Outmigration Dummy Variable (netmigdum), 2000, (1 if negative net migration)[b]	0.481	0.500	0.000	1.000
Median Age of Population (medage), 2000[b]	37.398	3.936	20.600	54.300
Median Age of Population Squared (medagesq), 2000[b]	1414.095	292.997	424.360	2948.490
Population Age 25 and Over with Bachelor's Degree or Higher (college), 2000 (% of total pop.)[b]	16.369	7.589	4.921	60.482
Population Age 25 and Over without High School Diploma (hsdeg), 2000 (% of total pop.)[b]	22.650	8.729	3.044	65.298
Ethnic Fragmentation Index (ethnic), 2000[b c]	0.214	0.169	0.005	0.707
Female Labor Force (female), 2000 (% of total labor force)[b]	45.850	2.123	23.051	54.053
Foreign Born Population (foreign), 2000 (% of total pop.)[b]	3.396	4.731	0.000	46.127

Data Sources and Explanations

[a] - Regional Economic Information System (REIS), Bureau of Economic Analysis (BEA)

[b] - U.S. Census Bureau, Census 2000

[c] - See text for further calculation details

[d] - Federal Deposit Insurance Corporation

[e] - United State Department of Agriculture (USDA), Economic Research Service (ERS), Creative Class County Codes

[f] - County Business Patterns

[g] - United State Department of Agriculture (USDA), Economic Research Service (ERS), Rural-Urban Continuum

* - Used in average entropy calculation and for sensitivity analysis

Methods

To perform the analysis, I use a state fixed-effects ordinary least squares (OLS) regression.[21] I include a dummy variable for each state to account for unobserved heterogeneity that might exist in policies and entrepreneurial climate across states. The regression model is as follows:

$$ED_{i,t+dt} = \beta_0 + \beta_1 X + \beta_2 FE + v_i$$

where on the RHS of the equation β_0 is a constant intercept term, β_1 is a vector of regression estimation coefficients for the *X* vector of exogenous variables described above, β_2 is a vector of regression estimation coefficients for the vector of state fixed effect dummy variables (*FE*), v_i is a random error term, *i* is a county index, *t* is initial time 2000, and *d* = 7.

A major concern in regression analysis is correlation among independent variables, or multicollinearity, which leads to potential heightening of standard errors in variables that are highly correlated. To test for multicollinearity, I create a correlation matrix for all variables included within the regression. The highest correlation among independent variables is .68 between the bohemian index and the percent of the population with a college degree or higher.[22] This falls within an acceptable range for a sample size of over 3,000. A full correlation analysis can be found in Appendix F. A second issue with OLS estimations is heteroscedasticity, which can cause bias in the variance and standard error of independent variables when present. I test for heteroscedasticity using White's test (1980) and find that it is present within the data. To correct for this, robust standard errors are used to account for the unequal variance in error terms (White, 1980). Finally, as with most regressions, there is a danger of finding biased and inconsistent

[21] The analysis is performed using STATA statistical analysis software.

[22] Median age and median age squared are highly correlated (.98), but the practice of including the squared term for age is common for estimating self-employment income equations (see Evans & Leighton, 1989).

results because of omission of variables. Based on previous research, I believe that the models incorporate all variables that are pertinent to entrepreneurial depth growth.

I begin by regressing (for all 3,047 counties) the change in entrepreneurial depth on the control variables, along with the measures for inmigration and outmigration volumes to assess the impact on self-employment income of having a large number of inflows or outflows in relation to county population. The subsequent regression introduces the entropy measure to test whether dispersed and diverse migrant networks are important for income growth. In the final four regressions, I seek to examine how brain drain from rural areas may be affecting self-employment income growth in those areas. The first two of the final regressions examine how migration volumes and network diversity are related to change in entrepreneurial depth in metro versus nonmetro areas, with the final two separating nonmetro counties into those that are adjacent to a metro county or not. By doing this, it is possible to see the effects of migration on the most rural of nonmetro areas. Results of the multivariate regressions are presented in the following chapter. All model presented below are appropriate, as F-statistics are significant at the 1% level.

Chapter 5
Empirical Results and Discussion

Multivariate Results: All U.S. Counties

Table 9 provides results from the regression estimations for all counties across the United States. The estimates take into account unobserved heterogeneity among states; however, individual state effect coefficients are not provided. Model 1 provides coefficient values without controlling for average entropy of the county, and Model 2 subsequently estimates change in entrepreneurial depth from 2000 to 2007 including the proxy for county migration stream diversity. The results coincide with previous theory in some aspects, but provide contrasting results in others.

Volumes of migration flows, as measured by inmigration and outmigration per capita, provide regression coefficients that are contrary to the expected outcomes. Outmigration per capita, originally hypothesized to be negatively and significantly associated with self-employment income growth, is not significant in the regression. This suggests that even if counties experience high "brain drain" of educated migrants, self-employment profits over the subsequent period are not affected. There may be a counteracting positive effect of information flow back to the home county through the network created by the migration.

Table 9: Multivariate Results: All Counties

	Results of Regression Equations: 3,047 U.S. Counties	
	Dependent Variable: Change in Entrepreneurial Depth, 2000-07	
	Model 1: Without Entropy	Model 2: Including Entropy
avgent	----	40.685***
	----	(3.72)
inmigPC	-33.028***	-30.577**
	(-2.72)	(-2.53)
outmigPC	-10.242	-6.565
	(-0.65)	(-0.42)
ED2000	-1.918***	-1.983***
	(-13.43)	(-13.54)
breadth	-0.305**	-0.222*
	(-2.31)	(-1.65)
risk	1.038***	1.038***
	(12.75)	(12.76)
deposits	0.254**	0.267**
	(1.96)	(2.09)
bsestab	0.225	0.040
	(0.71)	(0.13)
edestab	5.716***	4.763***
	(3.35)	(2.79)
bohemian	3.639*	2.775
	(1.65)	(1.24)
density	0.002***	0.002***
	(2.70)	(2.89)
cons	-0.235	-0.308
	(-0.98)	(-1.34)
nonmetro	0.651	0.003
	(0.47)	(0.00)
netmigdum	3.697***	3.481**
	(2.59)	(2.44)
medage	4.636***	5.657***
	(2.58)	(3.14)
medagesq	-0.062***	-0.075***
	(-2.64)	(-3.16)
college	-0.125	-0.149
	(-0.81)	(-0.96)
hsdeg	-0.072	0.046
	(-0.47)	(0.28)
ethnic	-3.504	-4.384
	(-0.60)	(-0.75)
female	-0.225	-0.344
	(-0.51)	(-0.80)
foreign	0.167	0.079
	(0.77)	(0.36)
constant	-36.358	-71.413
	(-0.85)	(-1.60)
R^2	0.341	0.345
N	3047	3047

Note: t-statistics in parenthesis — Significant at *10%, **5%, ***1%

Counties that experience large volumes of inflow in relation to its population also show lower growth in entrepreneurial depth over the time period in question. For each additional inmigrant per 1,000 population, entrepreneurial depth decreases by .03% for the time period. These results are also surprising, since it would be expected that counties that have high volume of inflows are receiving a higher educated population that is more likely to be able to realize potentially profitable business venture opportunities. One possible explanation for this may be that when migrants move to an area, they are able to recognize where current entrepreneurs are making profits and subsequently begin a competing business. This rationale is based upon the theory of Kirzner (1973, 1997), who suggests that many times outsiders to a situation have a "fresh perspective" and are able to notice opportunities which are not evident to people familiar with the situation (Glancey & McQuaid, 2000). Migrants moving to an area – outsiders - may be more perceptive to the fact that entrepreneurs are making a profit than people already living in the community. According to Kirzner, these migrants have the chance to follow in the footsteps of the prime mover and compete away their profits, which leads to lower entrepreneurial depth in counties with a higher rate of inmigration per capita.

After adding in average entropy measures, the model does not change significantly; however, it is clear that average entropy is positively and significantly correlated with growth in entrepreneurial depth. Counties with high values of entropy, or those with a diverse migration network, experience significantly more rapid growth than those with low entropy. An increase of average entropy of one tenth of a point (the index is a value on a 0 to 1 scale) leads to an increase in growth of entrepreneurial profits of over 4%. Counties with high average entropy have access to a wider breadth of information and novel knowledge via the dispersed network of migrants that have either left or come to the county during 1995 to 2000. The positively significant nature of this variable confirms the fact that diversity, and not just volume of flows, is an important characteristic of a county's migration network. This type of network centrality is indeed important in determining the creation of high value entrepreneurial endeavors, and subsequently is important to the growth of the local and regional economy.

Consistent with convergence literature, initial entrepreneurial depth and growth over the time period are inversely related. Self-employment incomes of counties grew faster in areas with initially lower levels of entrepreneurial depth in comparison to those with initially higher levels.

Results for county and economic controls are fairly consistent with theoretical insights. One of the strongest correlated variables is risk. Risk and change in entrepreneurial depth are positively associated with each other. Higher fluctuations in self-employment income led to higher change in entrepreneurial depth. Those in counties where self-employment was a riskier endeavor gained proportionately more in income per worker than in counties where entrepreneurship was less risky. Entrepreneurs are evidently able to perceive and capitalize upon the elevated riskiness in business ventures in counties characterized by dynamic self-employment returns. An overlapping result that supports Kirzner's theory is that entrepreneurial breadth is negatively (but not significantly) correlated to income growth. Higher number of entrepreneurs as a share of the total workforce leads to higher competition and lower profits among the self-employed. Bank deposits per capita are positively related to the dependent variable, confirming that higher initial levels of available capital improve the prospect of success for entrepreneurial ventures. Similarly, population density is positively associated with entrepreneurial depth growth, illustrating that access to a larger, denser market for products and services aids in building a more successful business. It may also signal that a higher concentration of population allows entrepreneurs greater choice in the labor market when choosing to hire. Higher population density provides a larger labor pool from which the entrepreneur can hire the most efficient workforce.

The number of support establishments is positively associated with entrepreneurial breadth growth; however, only the number of educational support establishments per capita is significant in the regression. A higher concentration per person of junior and technical colleges within a county positively influences self-employment return growth. Entrepreneurs are evidently taking advantage of the training opportunities that are available through these establishments. Self-employment profit growth is not related to the number of business support establishments, suggesting that entrepreneurs are failing to take advantage of opportunities to outsource basic business processes. If the self-employed are using business support establishments, the amount paid for services is sufficiently high enough to counteract positive gains from increased time to focus on core business development.

The employment mix within counties has no relationship with growth in entrepreneurial depth. Construction employment and the presence of a large creative class are not important in self-employment income growth. While both variables show signs in the expected direction, negative for construction employment and positive for the bohemian index, neither is significant at the 5% confidence level. Construction employment not being important in self-employment income growth may be due to the fact that incomes in the construction sector expanded during the time period in question due to historically high levels of new housing and commercial construction. The insignificance of the bohemian index indicates that even if there is a large concentration of artists and entertainers in a county, the inventiveness and creativity that comes with them is not being translated into high-value entrepreneurial ventures.

The final two county and economic controls, dummy variables for nonmetro and negative net migration counties, provide different results. The nonmetro dummy variable is not significant, showing that growth in entrepreneurial depth is statistically not different for nonmetro and metro counties when controlling for a variety of other variables. Nonmetro counties began with lower initial entrepreneurial depth than metro areas, but did not experience faster growth. This implies that even though convergence occurred for the entire U.S., the income gap between metro and nonmetro counties did not decrease, controlling for other variables. The net migration dummy, on the other hand, is positive and significant. This suggests that net outmigration counties experienced higher growth in self-employment incomes over the time period (nearly 3.5% higher) when compared to counties that gained population from migration. This again coincides with Kirzner's entrepreneurial arbitrage theory[23].

In terms of human capital and experience variables, results are mixed. Experience, as measured by median age and median age squared, has the expected relationship with entrepreneurial depth growth. The dependent variable and median age have an initially positive relationship, which peaks at around the age of 38 and subsequently remains negative. This result is remarkably consistent with the findings of Goetz & Shrestha (2009), who find a turning point at 37 years old. This confirms that there is a parabolic relationship between self-employment returns and experience, providing evidence for the fact that the older population is likely to become less concerned with entrepreneurial profits later in life. Variables that proxy human capital levels of the population, on the other hand, are not significant in determining change in profits over time. Having a bachelor's degree or lacking a high school education has no statistically significant relationship to growth. This may be a resulting factor of the "superstar effect" (Rosen, 1981), or the possibility that income growth for entrepreneurs was just as rapid in sectors of the economy that do not require a higher level of human capital (such as natural resource extraction or construction) as it was in those that do require higher human capital levels (such as financial services).

Finally, the demographic variables within the model are not significant in predicting change in entrepreneurial depth. The percent of the labor force that is female, the ethnic composition of the population, and the proportion of foreign born are not important to growth in

[23] I also estimated the model without the inclusion of the netmigdum variable to assess whether its inclusion had an influence on the other main independent migration variables. Results of this regression were not materially different than those including the variable, but the fit of the regression is better with the variable included.

self-employment profits. The ethnic fragmentation index and female labor force are negatively associated to entrepreneurial depth, in accordance with theory, and the proportion of foreign born is positively related. Again none of the demographic indicators is significant, suggesting that demographic composition of the county population was unimportant to growth in self-employment incomes.

Multivariate Results: Metro/Nonmetro County Segments

Within the next set of regressions, I utilize the same model of describing the growth in entrepreneurial depth. For these estimations, however, I first subset the U.S. counties into metro and nonmetro counties. The counties are then separated into more distinct nonmetro groups: those that are located adjacent to a metro county, and those that do not share a border with metro areas. Adjacency to a metro area is set by the Rural-Urban continuum code of the USDA ERS. Counties that are classified with a 4, 6, or 8 are all located next to a metro county. Those that are classified with a 5, 7, or 9 are not situated next to a metro area. I divide nonmetro counties into these groups to determine whether counties located close to a metro area behave differently than the most remote of counties. In other words, does being close to a larger urban agglomeration lead to different success outcomes for the self-employed? Results of all four regressions – one each for metro counties, nonmetro counties, nonmetro adjacent counties, and nonmetro non-adjacent counties – can be found in Table 10. I limit my discussion of results to those that deal with the most significant of the findings.

Table 10: Multivariate Results: Metro/Nonmetro Segments

	Results of Regression Equations: Metro/Nonmetro Segments *Dependent Variable: Change in Entrepreneurial Depth, 2000-07*			
	Metropolitan Counties	All Nonmetropolitan Counties	Nonmetro - Adjacent to Metro County	Nonmetro - Not Adjacent to a Metro County
avgent	4.422	63.882***	42.928*	75.484***
	(0.30)	(4.05)	(1.70)	(3.36)
inmigPC	-55.755***	-20.754	-26.206	-18.503
	(-3.08)	(-1.21)	(-1.30)	(-0.55)
outmigPC	42.233*	-24.852	-8.316	-31.644
	(1.67)	(-1.21)	(-0.31)	(-1.00)
ED2000	-1.573***	-2.761***	-3.158***	-2.511***
	(-7.90)	(-14.93)	(-12.40)	(-10.18)
breadth	-0.335	-0.280*	-0.630***	0.087
	(-1.34)	(-1.72)	(-2.79)	(0.37)
risk	0.805***	1.171***	1.222***	1.145***
	(6.13)	(10.85)	(7.56)	(8.84)
deposits	0.303	0.132	0.179	-0.050
	(1.26)	(1.10)	(1.18)	(-0.18)
edestab	3.789	3.417*	3.790	3.882*
	(1.01)	(1.78)	(1.25)	(1.66)
bsestab	1.245	-0.194	-0.436	-0.154
	(1.60)	(-0.61)	(-0.82)	(-0.37)
density	0.001**	-0.014	-0.016	0.027
	(2.56)	(-0.66)	(-0.69)	(0.48)
bohemian	2.760	3.376	-4.252	8.615**
	(0.69)	(1.19)	(-0.98)	(2.19)
cons	0.093	-0.295	0.088	-0.559
	(0.22)	(-1.06)	(0.25)	(-1.39)
netmigdum	4.344*	3.384*	1.616	3.880
	(1.82)	(1.82)	(0.77)	(1.17)
medage	5.486**	5.633**	2.840	7.632**
	(2.47)	(2.35)	(0.88)	(2.16)
medagesq	-0.075**	-0.076**	-0.034	-0.106**
	(-2.49)	(-2.42)	(-0.81)	(-2.27)
college	0.324	-0.753***	-0.566	-1.135***
	(1.29)	(-2.93)	(-1.58)	(-2.85)
hsdeg	0.581*	-0.284	-0.413	-0.362
	(1.65)	(-1.42)	(-1.57)	(-1.15)
ethnic	0.790	-12.137	0.977	-29.442**
	(0.08)	(-1.58)	(0.10)	(-2.31)
female	-0.841	-0.152	-0.546	0.194
	(-1.20)	(-0.29)	(-0.62)	(0.30)
foreign	-0.535*	0.300	0.136	0.457
	(-1.75)	(0.99)	(0.33)	(0.95)
constant	-60.491	-56.653	25.495	-97.124
	(-1.07)	(-0.97)	(0.29)	(-1.20)
R^2	0.366	.377	.448	.350
N	1043	2004	1042	962

Note: t-statistics in parenthesis — Significant at *10%, **5%, ***1%

There are some major differences in what determines the success of entrepreneurs in metro areas compared to nonmetro areas, especially in terms of migration volumes and networks. The first important difference between the two sets of counties is that entropy is not important for entrepreneurial success in metro areas, but is highly important in nonmetro counties. A highly dispersed inmigrant or outmigrant network has no relationship to growth in entrepreneurial depth in metro areas, but has a significant and positive effect for nonmetro areas. At the same time, the volume of in- and outflows show a significant relationship with growth in metro areas but not in nonmetro areas. The insignificant relationship between diversity of networks and entrepreneurial depth growth in metro areas may signify that there are alternative sources of novel knowledge that lead to the creation of high-value business ventures available to urban residents outside of a network of migrants. Interestingly, population density is also positively significant in metro areas, but not in nonmetro areas. This suggests that a higher population concentration in metro areas may be a source of alternative information. Entrepreneurs in metro areas may rely less on information passed through migrant networks and more on increased information due to knowledge spillovers in agglomeration economies (Rosenthal & Strange, 2004). Conversely, network entropy is highly significant in nonmetro areas. This suggests that entrepreneurs in rural counties that have diverse migration networks rely on the differentiated information flows for knowledge important to entrepreneurial development more so than their urban counterparts.

There are also significant differences among nonmetro counties that are adjacent to or not adjacent to a metro county. The most important difference is that migration network entropy is more important for entrepreneurial depth growth in the most rural of places (not-adjacent counties). In fact, for adjacent counties, average entropy is only significant at the 10% confidence level (t-statistic of 1.70). In non-adjacent counties, entropy is positively associated with entrepreneurial depth growth at the 1% confidence level. Both the correlation and magnitude of the entropy variable are stronger for non-adjacent counties. For each increase of .1

in the average entropy measure for not-adjacent counties, there is an average growth in entrepreneurial depth for the time period of 7.5%. For adjacent counties, there is just a 4.3% increase. Diversity of networks being less important in counties in close geographic proximity to larger cities suggests that there may be spillovers in entrepreneurial knowledge from urban areas. This finding supports the theory that geography and space are important in determining growth in innovative activities (Audretsch & Feldman, 2004), which in turn can lead to high-value entrepreneurial endeavors. The fact that non-adjacent county entropy is highly important in predicting entrepreneurial depth growth indicates that diverse information streams flowing from migrants within these counties are important sources of knowledge that can be translated to high-value business undertakings. Interestingly, the volume of in- or outmigration flows per capita are not important for growth in self-employment incomes in both segments of rural counties. This refutes Drabenstott's (2001) claim that rural brain drain is impeding the success of rural entrepreneurs. Additionally, Kirzner's arbitrage theory does not hold for either group of nonmetro counties. An increased number of outsiders (inmigrants per capita) is not associated with lower entrepreneurial profits in subsequent time periods.

In relation to the estimation results for control variables, the models between adjacent and non-adjacent counties also perform differently. For adjacent areas, there are very few significant variables, but this model predicts the change in entrepreneurial depth more effectively than any of the others, accounting for nearly 45% of the variance in entrepreneurial depth growth (R^2 of .448). For nonmetro non-adjacent areas, some of the control variables behave differently than in the other segments of the metro/nonmetro regressions. The proportion of college educated in the population actually has a negative, significant relationship with entrepreneurial depth growth. This surprising result may be due to the fact that the college educated may not enter into self-employment to the same extent as those lacking a bachelor's degree. High diversity in terms of race is negatively associated with self-employment income growth only in the most rural of areas, indicating the presence of the consumer discrimination that Meyer (1990) writes about. This makes sense in that the most rural of areas are traditionally the least diverse and most conservative.

Some of the control variables remain significant across all models regardless of metro or nonmetro status. Risk is positively related to entrepreneurial depth growth across all levels of rurality. Initial self-employment income is negatively and strongly correlated with growth in entrepreneurial depth, indicating that convergence is occurring within (but not across) different metro and nonmetro groupings. In nearly all models (except for that of adjacent metro counties), age shows an inverse U-shaped curve for entrepreneurial earnings growth.

Chapter 6
Summary and Conclusions

Main Findings

The most important outcome of this research is that entrepreneurs in counties with diverse or differentiated migration networks experience higher success than those in which migration networks are homogeneous. The principal explanation for this outcome has grounding in network theory: heterogeneous networks provide a wider breadth of knowledge flows to those within the network. Self-employed in counties that have high inmigration or outmigration entropy have access to differentiated knowledge and novel ideas that have been profitable in other locations. Those in high entropy counties also have more access to information on what business ventures were failures in other areas, which may be just as important as knowing which businesses were successful.

Furthermore, the importance of diversity in migration networks differs by level of rurality. By examining results from regressions separating counties on the basis of the Rural-Urban continuum code, I find that diversity within the migrant network is not important for self-employment income growth in metro counties. Contrastingly, a highly dispersed network is important for the growth of entrepreneurial profits in nonmetro areas, especially so for the most rural of counties that are not adjacent to a metro area. Proximity to a metro county lowers the importance of a diverse migration network, suggesting that knowledge spillovers within urban areas or from urban areas to adjacent counties are important sources of knowledge for the growth of successful entrepreneurial ventures.

Secondarily, I find that outmigration volumes are not related to growth in self-employment incomes and high inmigration volumes are related to lower self-employment profits. These outcomes are contrary to what was originally hypothesized. This refutes previous theory that the rural brain drain impedes entrepreneurial success in rural areas (Drabenstott, 2001). This outcome also supports Kirzner's entrepreneurial arbitrage theory in which outsiders (migrants) enter in a new situation, recognize where profits are being made, and subsequently compete them away by beginning a rival venture.

Policy Implications

Counties with diverse migration streams experienced more rapid growth in self-employment incomes during the 2000s, indicating that information that can be translated into assistance for local entrepreneurs travels through the network streams. Furthering the ease of transmission or reducing the costs of transmission of information may help to increase future knowledge transfers. This can be done through investment in infrastructure that makes it easier for communication to occur between migrants and their home counties or that facilities movement of migrants to areas that are non-traditional destination locations. Increased improvements to broadband internet access and improved telecommunications networks would allow migrants to converse more easily with those still in the home community. Investment in communications systems would benefit the most rural of areas due to the fact that these areas are the most likely to lack access to broadband internet or services of equivalent speed. Investment in transportation infrastructure that allows migrants to move to a wider variety of areas would also aid in stimulating entrepreneurial success, again especially so for the most rural of areas that tend to have lower average entropy values (see Figure 11).

Higher investment in entrepreneurial programs in areas with low entropy values may also help to boost self-employment profits. In areas lacking high entropy, there is less diverse information flowing via migration networks to or from the county. Creation or promotion of programs that provide information to train the self-employed (and/or those who may potentially start a business) in creating high-value businesses may be an effective alternative where an expansive network of migrants is lacking. Programs to aid entrepreneurs could include valuable information about what business ventures have been successful in geographically diverse areas. Contrarily, information about business failures in other locations would be another important piece of knowledge to share.

Future Research

There is a great deal of opportunity for future research. Empirical studies merging migration network theory and entrepreneurial outcomes are basically nonexistent. This study serves as an exploratory analysis of the topic at hand. Further works need to be completed in this area to verify and solidify the relationships found within this paper. One major drawback of this type of research is that, because data are cross-sectional, it is not possible to know if having diverse migration patterns causes self-employment income growth. It is only possible to say that there is a positive relationship between the two. The correlations within this research are simply associations; however, the results here are an important starting point for further research. Historical data from many Census years could be used in a panel analysis in order to gain a greater understanding of causal relationships instead of just cross-sectional associations. Additional measures of entrepreneurial success, such as firm survival rate, can serve as a substitute for the growth in entrepreneurial depth as a dependent variable. Other data sources on internal U.S. migration, such as the American Community Survey (ACS) or the Internal Revenue

Service (IRS) can be used to create entropy measures for different time periods in order to test the results obtained here. More advanced statistical analyses such as geographically weighted regressions or spatial econometric analysis are other useful tools that could be used to assess the significance of the relationship between migration networks and entrepreneurial success in a spatial context.

A more qualitative approach to analyzing the relationship between these two economically important phenomena would provide further richness. Studies that survey entrepreneurs to obtain information about their personal migration history or the contacts that they maintain with others who have migrated can provide valuable micro-level information on how the self-employed are using information gained through migration networks. Questions about where the respondents have migrated to/from, who they maintain linkages with in home areas, and contact with those have moved to specific destinations would allow for analysis of the diversity of network flows and the extent of the differentiation of knowledge that is available. As stated earlier, this area of research is virtually unexplored; therefore, it provides ample opportunity for contribution to knowledge of how migration networks influence entrepreneurship at the micro and macro level.

Explicitly stating questions about remaining uncertainty can help to direct future research. Some specific questions that remain are:

- Does the positive relationship between self-employment income and migration networks hold over different time periods?
- Can the same relationship be supported at the micro level? Do entrepreneurs' reported migration experiences coincide with aggregate, county level data?
- Can implementation of policies that improve knowledge and business information for the self-employed in counties with less diverse migration networks affect their income?

- Does the diversity of migration networks show the same positive association with alternative proxies for entrepreneurial success such as firm births and deaths?

Understanding how migration networks affect self-employment income growth is just one instance of how diversity in information flows affects an important aspect of economic development. Further exploration of how migration network entropy may affect other determinants of economic growth would assist in helping to understand how communities and regions grow. Further knowledge of how networks affect regional economic development overall is key to expanding our understanding of what influences economic success and failure across the United States.

Appendix A – Rural-Urban Continuum Description

2003 Rural-Urban Continuum Codes	
Code	**Description**
Metro counties:	
1	Counties in metro areas of 1 million population or more
2	Counties in metro areas of 250,000 to 1 million population
3	Counties in metro areas of fewer than 250,000 population
Nonmetro counties:	
4	Urban population of 20,000 or more, adjacent to a metro area
5	Urban population of 20,000 or more, not adjacent to a metro area
6	Urban population of 2,500 to 19,999, adjacent to a metro area
7	Urban population of 2,500 to 19,999, not adjacent to a metro area
8	Completely rural or less than 2,500 urban population, adjacent to a metro area
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area
Source: USDA ERS, 2004	

Appendix B – Self-employment Satisfaction

Job Satisfaction
Employees vs. Self-employed

	Employees	Self-employed
Very dissatisfied	3.8%	2.0%
A little dissatisfied	10.3%	5.8%
Moderately satisfied	39.9%	29.8%
Completely satisfied	45.9%	62.5%

Source: Blanchflower, 2004; Using the U.S. General Social Survey

Appendix C – Alternative Outmigration and Inmigration Entropy Regressions

Results of Separate Outmigration and Inmigration Entropy Regression Equations

	Dependent Variable: Change in Entrepreneurial Depth, 2000-07	
	Outmigration Entropy Coefficient Results	Inmigration Entropy Coefficient Results
outentropy	31.906***	
	(3.03)	
inentropy		39.192***
		(4.03)
inmigPC	-31.314***	-29.856**
	(-2.59)	(-2.48)
outmigPC	-7.097	-9.144
	(-0.46)	(-0.60)
ED2000	-1.975***	-1.989***
	(-13.46)	(-13.72)
breadth	-0.230*	-0.215
	(-1.73)	(-1.63)
risk	1.036***	1.037***
	(12.71)	(12.76)
deposits	0.272**	0.270**
	(2.08)	(2.12)
edestab	4.980***	4.661***
	(2.93)	(2.73)
bsestab	0.079	0.047
	(0.26)	(0.16)
density	0.002***	0.002***
	(2.81)	(2.83)
bohemian	3.131	2.762
	(1.40)	(1.23)
cons	-0.296	-0.278
	(-1.29)	(-1.20)
netmigdum	3.487**	3.406**
	(2.44)	(2.38)
nonmetro	0.395	-0.224
	(0.28)	(-0.16)
medage	5.356***	5.651***
	(2.99)	(3.15)
medagesq	-0.070***	-0.075***
	(-2.99)	(-3.19)
college	-0.135	-0.153
	(-0.87)	(-1.00)
hsdeg	0.024	0.02
	(0.15)	(0.13)
ethnic	-4.452	-4.049
	(-0.76)	(-0.70)
female	-0.383	-0.413
	(-0.93)	(-1.00)
foreign	0.068	0.063
	(0.32)	(0.30)
constant	-64.026	-69.931
	(-1.43)	(-1.58)
R^2	0.344	0.346
N	3047	3047
Note: t-statistics in parenthesis		Significant at *10%, **5%, ***1%

Appendix D – Examples of Entropy Measure Construction

Construction of Outmigration Entropy: Storey County, Nevada

Origin: Storey County, Nevada

Destination County	Out Flow (m_{ij})	p_{ij}	$\log_2 p_{ij}$	$p_{ij} * \log_2 p_{ij}$
Alameda County, California	17	0.031657	-4.98132	-0.1576953
Amador County, California	2	0.003724	-8.06878	-0.0300513
Calaveras County, California	16	0.029795	-5.06878	-0.1510251
Los Angeles County, California	17	0.031657	-4.98132	-0.1576953
San Bernardino County, California	4	0.007449	-7.06878	-0.0526538
San Francisco County, California	9	0.01676	-5.89885	-0.0988635
Idaho County, Idaho	17	0.031657	-4.98132	-0.1576953
Webster County, Iowa	4	0.007449	-7.06878	-0.0526538
Webster County, Missouri	3	0.005587	-7.48382	-0.041809
Gallatin County, Montana	8	0.014898	-6.06878	-0.0904101
Eureka County, Nevada	5	0.009311	-6.74685	-0.0628198
Humboldt County, Nevada	38	0.070764	-3.82085	-0.2703768
Lyon County, Nevada	11	0.020484	-5.60935	-0.1149028
Washoe County, Nevada	255	0.47486	-1.07442	-0.5102017
Carson City, Nevada	57	0.106145	-3.23589	-0.3434742
Richland County, Ohio	6	0.011173	-6.48382	-0.0724449
Oklahoma County, Oklahoma	9	0.01676	-5.89885	-0.0988635
Utah County, Utah	53	0.098696	-3.34086	-0.3297308
Clark County, Washington	6	0.011173	-6.48382	-0.0724449
Total Number of Outmigrants ($\sum m_{ij}$)	537			

$-\sum(pij * \log2pij)$ =	2.865811858
$\log_2 N$ =	11.57317378
ε =	0.247625406

Construction of Outmigration Entropy: Billings County, North Dakota

Origin: Billings County, North Dakota

Destination County	Out Flow (m_{ij})	p_{ij}	$\log_2 p_{ij}$	$p_{ij} * \log_2 p_{ij}$
Mariposa County, California	8	0.04188	-4.57743	-0.1917248
San Bernardino County, California	7	0.03665	-4.77007	-0.1748195
Fremont County, Idaho	5	0.02618	-5.2555	-0.1375786
Saline County, Kansas	4	0.02094	-5.57743	-0.1168048
Adams County, North Dakota	2	0.01047	-6.57743	-0.0688736
Bowman County, North Dakota	8	0.04188	-4.57743	-0.1917248
Burke County, North Dakota	2	0.01047	-6.57743	-0.0688736
Cass County, North Dakota	5	0.02618	-5.2555	-0.1375786
Golden Valley County, North Dakota	2	0.01047	-6.57743	-0.0688736
Logan County, North Dakota	1	0.00524	-7.57743	-0.0396724
Richland County, North Dakota	10	0.05236	-4.2555	-0.2228011
Stark County, North Dakota	95	0.49738	-1.00757	-0.501149
Ward County, North Dakota	3	0.01571	-5.99247	-0.0941225
Wells County, North Dakota	2	0.01047	-6.57743	-0.0688736
Jefferson County, Oregon	7	0.03665	-4.77007	-0.1748195
Corson County, South Dakota	22	0.11518	-3.118	-0.359141
Meade County, South Dakota	8	0.04188	-4.57743	-0.1917248
Total Number of Outmigrants ($\sum m_{ij}$)	191			
$-\sum(pij * \log2pij) =$	2.809155547			
$\log_2 N =$	11.57317378			
$\varepsilon =$	0.24272992			

Appendix E – Age of Migrants from the CPS

Domestic Gross County-to-County Migration Figures
1990-1995 (in thousands)

Age Cohorts	*Total County-to-County Movers*	*Different County, Same State*	*Percent of Total*	*Diff State, Same Division*	*Percent of Total*	*Different Division, Same Region*	*Percent of Total*	*Different Region*	*Percent of Total*
Total +5 Years	40,998	21,311	52.0%	5,570	13.6%	3,814	9.3%	10,303	25.1%
5-9 years	3,862	1,987	51.5%	520	13.5%	366	9.5%	989	25.6%
10-14 years	3,142	1,547	49.2%	450	14.3%	290	9.2%	855	27.2%
15-17 years	1,475	747	50.6%	228	15.5%	137	9.3%	363	24.6%
18-19 years	975	512	52.5%	118	12.1%	92	9.4%	253	25.9%
20-24 years	4,645	2,493	53.7%	564	12.1%	397	8.5%	1,191	25.6%
25-29 years	5,989	3,060	51.1%	818	13.7%	584	9.8%	1,527	25.5%
30-34 years	5,464	2,952	54.0%	725	13.3%	487	8.9%	1,300	23.8%
35-39 years	4,301	2,254	52.4%	610	14.2%	402	9.3%	1,035	24.1%
40-44 years	3,182	1,639	51.5%	448	14.1%	251	7.9%	844	26.5%
45-49 years	2,176	1,153	53.0%	289	13.3%	199	9.1%	535	24.6%
50-54 years	1,583	801	50.6%	194	12.3%	189	11.9%	399	25.2%
55-59 years	1,111	581	52.3%	180	16.2%	113	10.2%	237	21.3%
60+ years	3,087	1,584	51.3%	423	13.7%	307	9.9%	773	25.0%

Source: U.S. Census Bureau, Current Population Survey (2000).
http://www.census.gov/population/www/socdemo/migrate/p23-200.html

Appendix F – Correlation Analysis

Correlation Analysis

	avgent	*inmigPC*	*outmigPC*	*medage*	*medagesq*	*college*	*hsdeg*	*female*	*foreign*	*edestab*	*bsestab*	*risk*	*ED2000*
avgent	1												
inmigPC	0.0971131	1											
outmigPC	0.0640747	0.4962543	1										
medage	-0.2418386	-0.1449801	-0.0985312	1									
medagesq	-0.2349802	-0.1240937	-0.0715097	0.995561	1								
college	0.4584244	0.3931551	0.3829563	-0.1870893	-0.1741838	1							
hsdeg	-0.2896739	-0.3249674	-0.3491361	-0.1179399	-0.1246648	-0.651536	1						
female	0.2180312	-0.1459658	-0.2404539	-0.0238707	-0.0292902	0.1427373	-0.0713601	1					
foreign	0.2446695	0.0893428	0.1927393	-0.3138006	-0.3014841	0.3495211	0.0475506	-0.1659821	1				
edestab	0.3668888	0.1094994	0.0892705	-0.1508147	-0.1485017	0.4176784	-0.2552394	0.1586412	0.1818137	1			
bsestab	0.417271	0.0738455	0.1309168	-0.0550714	-0.0578621	0.4638816	-0.3548285	0.1733016	0.1481834	0.2627599	1		
risk	0.0442022	0.2114114	0.0510128	-0.1340471	-0.1343811	0.1848188	-0.0670023	0.0049735	0.0779683	0.054632	0.0553505	1	
ED2000	0.4055452	-0.0760093	-0.03385	-0.171108	-0.1797068	0.3717536	-0.1107258	0.1569474	0.3678285	0.2874573	0.3273788	0.0578167	1
breadth	-0.3242885	0.2517966	0.0990711	0.3782332	0.3824279	-0.0580161	-0.0881018	-0.2227345	-0.1212656	-0.1353947	-0.1473392	0.2332656	-0.3143473
bohemian	0.3925685	0.2915027	0.2189911	-0.0937557	-0.0851887	0.679944	-0.4230402	0.100067	0.2979507	0.3329528	0.3432832	0.1203707	0.3596809
cons	0.0838239	0.3109019	-0.0739006	-0.0032817	-0.0098727	0.1260542	-0.1231833	-0.021845	-0.0368086	0.0592359	0.0696904	0.2256863	-0.0055569
density	0.0855513	-0.0376461	0.0162533	-0.0722887	-0.0737872	0.1889537	-0.0206208	0.0990284	0.3467388	0.1304387	0.1026115	0.0151848	0.3867226
deposits	0.0700716	-0.1589827	0.0137258	0.203504	0.2029204	0.1859112	-0.2000028	0.0664701	0.0713874	0.0774698	0.200446	-0.0894891	0.2482129
nonmetro	-0.2161001	-0.2214413	0.0186145	0.2856026	0.2926791	-0.3671938	0.2329485	-0.1130046	-0.2255078	-0.2366589	-0.1502679	-0.2113195	-0.2812752
netmigdum	-0.022805	-0.519443	0.1608453	0.081431	0.081887	-0.0958194	0.0269491	-0.0116113	0.0656231	-0.0365182	0.0034634	-0.1687675	0.0750577

Correlation Analysis, Cont.

breadth	*bohemian*	*cons*	*density*	*deposits*	*nonmetro*	*netmigdum*
1						
0.0268056	1					
0.3378549	0.2183899	1				
-0.0910919	0.2902694	-0.025594	1			
-0.0492144	0.1635054	-0.1166335	0.2736428	1		
0.0734109	-0.2709519	-0.2653189	-0.1505636	0.0928298	1	
-0.1747783	-0.1196417	-0.299875	0.0561888	0.1699426	0.1949419	1

References

Acs, Z.J. & Armington, C. (2006). *Entrepreneurship, geography, and American economic growth.* New York: Cambridge University Press.

Aldrich, H.E. & Zimmer, C. (1986). Entrepreneurship through social networks. In D. L. Sexton & R. W. Smilor (Eds.). *The art and science of entrepreneurship* (pp. 3–23). New York: Ballinger.

Alesina, A., Baqir, R. & Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics, 114,* 1243-1284.

Artz, G. (2003). Rural area brain drain: Is it a reality? *Choices, 4th Quarter,* 11-16.

Audretsch, D.B. & Feldman, M.P. (2004). Knowledge spillovers and the geography of innovation. In V. Henderson & J.F. Thisse (Eds.) *Handbook of urban and regional economics.* (pp. 2713-2739). Elsiever: retrieved Feb 15, 2011, from http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B7P64-4FFPH84-2R&_user=209810&_coverDate=12%2F31%2F2004&_rdoc=1&_fmt=high&_orig=search&_origin=search&_sort=d&_docanchor=&view=c&_acct=C000014439&_version=1&_urlVersion

Audretsch, D.B. & Keilbach, M. (2004). Entrepreneurship capital and economic performance. *Regional Studies, 38,* 949-959.

Audretsch, D.B. & Thurik, A.R. (2000). Capitalism and democracy in the 21st century: From the managed to the entrepreneurial economy, *Journal of Evolutionary Economics, 10,* 17-34.

Barabási, A.L. (2002). *Linked: The new science of networks.* Cambridge: Perseus.

Bates, T. (1990). Entrepreneur human capital inputs and small business longevity. *The Review of Economics and Statistics, 72,* 551-559.

Bates, T. (1997). *Race, self-employment, and upward mobility: an illusive American dream.* Baltimore, MD: Johns Hopkins University Press.

Bednarzik, R.W. (2000, July). *The role of entrepreneurship in U.S. and European job growth.* (U.S. Bureau of Labor Statistics, Monthly Labor Review). Washington, D.C.: Government Printing Office.

Benz, M. and Frey, B. S. (2008). Being independent is a great thing: Subjective evaluations of self-employment and hierarchy. *Economica, 75,* 362–383.

Beyers, W.B. & Nelson, P.B. (2000). Contemporary development forces in the nonmetropolitan west: new insights from rapidly growing communities. *Journal of Rural Studies,* 16, 459-474.

Birch, D.L. (1979). *The job generating process.* Cambridge, MA: MIT Programme on Neighborhood and Regional Change.

Birch, D.L. (1981). Who creates jobs? *Public Interest, Fall,* 3-14.

Blanchflower, D.G. (2004). Self-employment: More may not be better (NBER working paper No.10286). *NBER Working Paper Series*. Cambridge, MA: National Bureau of Economic Research.

Blanchflower, D.G. & Oswald, A.J. (1992). Entrepreneurship, happiness, and supernormal returns: Evidence from Britain and the US (NBER working paper No.4228). *NBER Working Paper Series*. Cambridge, MA: National Bureau of Economic Research.

Blanchflower, D.G. & Oswald, A.J. (1998). What makes an entrepreneur? *Journal of Labor Economics, 16,* 26-60.

Blau, D.M. (1987). A time series analysis of self-employment in the United States. *The Journal of Political Economy, 95,* 445-467.

Borgatti, S., Mehta, D.J., Brass, D.J., & Lanianca, G. (2009). Network analysis in the social sciences. *Science, 323,* 892-895.

Borjas, G.J. (1990). *Friends or strangers, the impact of immigrants on the U.S. economy.* New York: Basic Books.

Borjas, G.J. & Bronars, S.J. (1989). Discrimination and self-employment. *The Journal of Political Economy, 97,* 581-605.

Borjas, G.J, Bronars, S.G. & Trejo S.J. (1992). Self-selection and internal migration in the United States. *Journal of Urban Economics, 32,* 159-185.

Bowles, S. (1970). Migration as investment: Empirical tests of the human investment approach to geographical mobility. *Review of Economics and Statistics, 52,* 356-365.

Brüdler, J. & Preisendörfer, P. (1998). Network support and the success of newly founded businesses. *Small Business Economics, 10,* 213-225.

Budig, M.J. (2006). Gender, Self-employment, and earnings: The interlocking structures of family and professional status. *Gender & Society, 60,* 725-753.

Buttner, E.H. & Moore, D.P. (1997). Women's organizational exodus to entrepreneurship: Self-reported motivations and correlates with success. *Journal of Small Business Management, 37,* 34-46.

Camarota, S. (2000). *Reconsidering immigrant entrepreneurship: An examination of self-employment among natives and the foreign-born.* Washington, DC: Center for Immigration Studies.

Clayton, R. & Spletzer, J. (2006). *A profile of US data source on entrepreneurship.* Washington, DC: Government Printing Office. Retrieved Jan 31, 2011, from http://www.oecd.org/dataoecd/24/46/37822254.pdf

Coleman, J., Katz, E., & Menzel, H. (1966). The diffusion of an innovation among physicians. *Sociometry, 20,* 253-270.

Cooper, A.C., Gimeno-Gascon, F.J., & Woo, C.Y. (1994). Initial human and financial capital as predictors of new venture performance. *Journal of Business Venturing, 9,* 371-395.

Courchene, T.J. (1970). Inter-provincial migration and economic adjustment. *Canadian Journal of Economics, 3,* 550-76.

Deller, S.C. & Goetz, S.J. (2009). A brief history of economic development policy and thought. In S.J. Goetz, S.C. Deller, & T.R. Harris (Eds.). *Targeting Regional Economic Development* (pp. 17-34). UK: Routledge Studies in Global Competition.

Domina, T. (2006). What clean break?: Education and nonmetropolitan migration patterns 1989-2004. *Rural Sociology, 71,* 373-391.

Douglas, E. & Shepherd, D. (2002). Self-employment as a career choice: Attitudes, entrepreneurial intentions, and utility maximization. *Entrepreneurial Theory and Practice, 26*, 81-90.

Drabenstott, M. (2001). New policies for a new rural America. *International Regional Science Review, 24*, 3-15.

Dubini, P. & Aldrich, H. (1991). Personal and extended networks are important to the entrepreneurial process. *Journal of Business Venturing, 6,* 305-313.

Duchesneau, D.A. & Gartner, W.B. (1990). A profile of new venture success and failure in an emerging industry. *Journal of Business Venturing*, 5, 297-312.

Duncan, G.J. (1976). Earnings functions and nonpecuniary benefits. *The Journal of Human Resources, 11*, 462-483.

Duncan, G.J. & Stafford, F.P. (1980). Do union members receive compensating wage differentials? *The American Economic Review, 70,* 355–371.

Dunn, T. & Holtz-Eakin, D. (2000). Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links. *Journal of Labor Economics,* 18, 282-305.

Durlauf, S.N., Johnson, P.A. & Temple, J.R.W. (2005). Growth econometrics. In P. Aghion & S. Durlauf (Eds.). *Handbook of economic growth* (pp. 555–677). Burlington: Elsevier.

Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. *Science, 328,* 1029-1031.

Evans, D.S. & Leighton, L.S. (1989). Some empirical aspects of entrepreneurship. *The American Economic Review, 79,* 519-535.

Federal Deposit Insurance Corporation (FDIC). (2004). *Summary of deposits.* Retrieved Oct 23, 2010, from http://www2.fdic.gov/sod/dynaDownload.asp?barItem=6

Findeis, J.L. & Jensen, L. (1998). Employment opportunities in rural areas: Implications for poverty in a changing policy environment. *American Journal of Agricultural Economics, 80,* 1000-1007.

Florida, R. (2002a). Bohemia and economic geography. *Journal of Economic Geography*, 2, 55-71.

Florida, R. (2002b). The economic geography of talent. *Annals of the association of American geographers, 92,* 743-755.

Folger, J.K. & Nam, C.B. (1967). *Education of the American population.* (U.S. Census Bureau 1960 Census Monogram) Washington, D.C.: Government Printing Office.

Fortunato, M.W-P. & Alter, T.R. (2011). The individual-institutional-opportunity nexus: An integrated framework for analyzing entrepreneurship development. *Entrepreneurship Research Journal, 1*, Article 6.

Franklin, R.S. (2003a). *Domestic migration across regions, divisions, and states: 1995 to 2000.* (U.S. Census Bureau Census 2000 Special Reports CENSR-7). Washington, D.C.: Government Printing Office.

Franklin, R.S. (2003b). *Migration of the young, single, and college educated: 1995 to 2000.* (U.S. Census Bureau Census 2000 Special Reports CENSR-12). Washington, D.C.: Government Printing Office.

Freeman, L.C. (1979). Centrality in networks: I. Conceptual clarification. *Social Networks, 1,* 215–239.

Fuchs, V.R. (1982). Self-employment and labor force participation of older males. *Journal of Human Resources, 17,* 339-357.

Gaillard, J. & Gaillard, A.M. (1998). The international circulation of scientists and technologists: A win-lose or win-win situation? *Science Communication, 20,* 106-115.

Gartner, W.B. (1990). What are we talking about when we talk about entrepreneurship? *Journal of Business Venturing, 5,* 15-28.

Georgellis, Y. & Wall, H.J. (2005). Gender differences in self-employment. *International Review of Applied Economics, 19,* 321-342.

Glaeser, E.L. (2001). Cities, information, and economic growth. *Cityscape, 1,* 9-47.

Glancey, K.S. & McQuiad, R.W. (2000). *Entrepreneurial economics.* New York: St. Martin's Press, Inc.

Goetz, S.J., Han, Y., Findeis, J.L., & Brasier, K.J. (2010). U.S. commuting networks and economic growth: Measurement and implications for spatial policy. *Growth and Change, 41*, 276-302.

Goetz, S.J. & Rupasingha, A. (2003). The returns on higher education: Estimates for the 48 contiguous states. *Economic Development Quarterly, 17,* 337-351.

Goetz, S. J. & Rupasingha, A. (2009). Determinants of growth in non-farm proprietor densities in the US, 1990-2000. *Small Business Econ, 32,* 425-438.

Goetz, S.J. & Shrestha, S.S. (2009). Explaining self-employment success and failure: Wal-Mart versus Starbucks, or Schumpeter versus Putnam. *Social Science Quarterly, 90,* 22-38.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory, 1,* 201–233.

Greenwood, M.J. (1975). Research on internal migration in the United States: A survey. *Journal of Economic Literature, 12,* 397-433.

Hamilton, B. H. (2000). Does entrepreneurship pay? An empirical analysis of the returns of self-employment. *The Journal of Political Economy, 108,* 604-631.

Hanneman, R.A. & Riddle, M. (2005). *Introduction to social network methods.* Riverside, CA: University of California, Riverside. Retrieved Jan 25, 2010, from http://faculty.ucr.edu/~hanneman/

Haug, S. (2008). Migration networks and migration decision-making. *Journal of Ethnic and Migration Studies, 34,* 585-605.

Higgins, M.J., Levy, D., & Young, A.T. (2006). Growth and convergence across the United States: Evidence from county-level data. *Review of Economics and Statistics, 88,* 671-681.

Hoang, H. & Antoncic, B. (2003). Network based research in entrepreneurship: A critical review. *Journal of Business Venturing, 18,* 165-187.

Holtz-Eakin, D., Joulfaian, D., & Rosen, H.S. (1994). Entrepreneurial decisions and liquidity constraints. *The RAND Journal of Economics, 25,* 334-347.

Hundley, G. (2000). Male/female earnings differentials in self-employment: The effects of marriage, children, and the household division of labor. *Industrial and Labor Relations Review, 54,* 95-114.

Hunter, L.M. (1998). The association between environmental risk and internal migration flows. *Population and Environment, 19,* 247-277.

Irwin, E.G., Isserman, A.M., Kilkenny, M., & Partridge, M.D. (2010). A century of research on rural development and regional issues. *American Journal of Agricultural Economics, 92,* 522-553.

Iversen, J., Jorgensen, R. & Malchow-Moller, N. (2008, January). Measuring Entrepreneurship. *Foundations and Trends in Entrepreneurship*. Retrieved November 22, 2010, from http://www.entrepreneur.com/tradejournals/article/171539784.html

Johnson, J.M. & Regets, M.C. (1998). International mobility of scientists and engineers to the U.S.: Brain drain or brain circulation? *National Science Foundation Issue Brief,* 98-316. Retrieved Feb 4, 2011, from http://www.nsf.gov/sbe/srs/issuebrf/sib98316.htm

Johnson, K.M. (2001). The continuing incidence of natural decrease in American counties. *Rural Sociology, 76,* 74-100.

Kihlstrom, R.E. & Laffont, J.J. (1979). A general equilibrium entrepreneurial theory of firm formation based on risk aversion, *Journal of Political Economy, 87,* 719-749.

Kilkenny, M.R. & Johnston, S.R. (2007). *Rural development policy.* Retrieved Jan 29, 2011, from http://www.cabnr.unr.edu/kilkenny/Rural%20Development%20Policy%20for%20AEI.pdf

Kirzner, I.M. (1973). *Competition and entrepreneurship.* Chicago: University of Chicago Press.

Kirzner, I.M. (1997). Entrepreneurial discovery and the competitive market process: An Austrian approach. *Journal of Economic Literature, 35,* 60-85.

Knight, F.H. (1921). *Risk, uncertainty, and profit.* New York: Houghton Mifflin.

Kusmin, L., Gibbs, R. & Parker, T. (2008). Education’s role in the metro-nonmetro earnings divide. *Amber Waves, 6,* 30-35.

Lee, S.Y., Florida, R., & Acs, Z.J. (2004). Creativity and entrepreneurship: A regional analysis of new firm formation. *Regional Studies, 38,* 879-891.

Lobo, J. & Smole, D.P. (2002). Stratification and spatial segregation of human capital as determinants of metropolitan productivity in the United States. *Urban Studies, 39,* 529-547.

Low, S.A. (2004, September). Regional asset indicators: Entrepreneurship breadth and depth. *The Main Street Economist.* Retrieved November 22, 2010, from http://www.kansascityfed.org/Publicat/mse/MSE_0904.pdf

Low, S.A., Henderson, J., and Weiler, S. (2005). Gauging a region's entrepreneurial potential. *Economic Review, Federal Reserve Bank of Kansas City,* 61-89.

McGranahan, D., Cromartie, J, & Wojan, T. (2010). *Nonmetropolitan outmigration counties: Some are poor, many are prosperous.* (U.S. Department of Agriculture, ERS. Economic Research Report 107). Washington, D.C.: Government Printing Office.

McGranahan, D. & Wojan, T. (2007). *Creative class county codes.* Retrieved Nov 11, 2010, from http://www.ers.usda.gov/Data/CreativeClassCodes/

McKernan, S. & Salzman, H. (2008). *Self-employment and economic mobility*. Washington, DC: The Urban Institute.

Meyer, B.D. (1990). Why are there so few black entrepreneurs? (NBER working paper No.3537). *NBER Working Paper Series*. Cambridge, MA: National Bureau of Economic Research.

Meyer, J.B. (2001). Network approach versus brain drain: Lessons from diaspora. *International Migration, 39,* 91-110.

Millimet, D.L. & Osang, T. (2007). Do state borders matter for U.S. intranational trade? The role of history and internal migration. *Canadian Journal of Economics, 40,* 93-126.

Morin, R. (2009). *Take this job and love it.* Washington, DC: Pew Research Center. Retrieved January 9, 2011, from http://pewsocialtrends.org/2009/09/17/take-this-job-and-love-it/

Nelson, P. (1959). Migration, real income, and information. *Journal of Regional Science, 1,* 43-74.

Nogle, J.M. (1994). The systems approach to international migration: An application of network analysis methods. *International Migration, 32,* 329-342.

O'Farrell, P.N. (1986). Entrepreneurship and regional development: Some conceptual issues. *Regional Studies, 20,* 565-574.

Ogle, R. (2007). *Smart world: Breakthrough and creativity and the new science of ideas.* Boston: Harvard Business School Publishing.

Page, S. (2007). *The difference: How the power of diversity creates better groups, firms, schools, and societies.* Princeton, NJ: Princeton University Press.

Parker, S.C. (1996) A time series model of self-employment under uncertainty. *Economica*, 63, 459-475.

Portes, A. & Zhou, M. (1996). Self-employment and the earnings of immigrants. *American Sociological Review, 61,* 219-230.

Rauch, J.E. (1993). Productivity gains from geographic concentrations of human capital: Evidence from cities. *Journal of Urban Economics, 34*, 380-400.

Rees, H. & Shah, A. (1986). An empirical analysis of self-employment in the U.K. *Journal of Applied Econometrics, 1,* 95-108.

Robinson, P.B. & Sexton, E.A. (1994). The effect of education and experience on self-employment success. *Journal of Business Venturing, 9*, 141-156.

Rosen, S. (1981). The economics of superstars. *The American Economic Review, 71*, 845-858.

Rosenthal, S.S. & Strange, W.C. (2004). Evidence on the nature and sources of agglomeration economies. In V. Henderson & J.F. Thisse (Eds.) *Handbook of urban and regional economics.* (pp. 2119-2171). Elsiever: retrieved Feb 15, 2011, from http://www.sciencedirect.com/science/article/B7P64-4FFPH84-29/2/5406781bb197c83b1589165139006bc9

Rosti, L. & Chelli, F. (2005). Gender discrimination, entrepreneurial talent, and self-employment. *Small Business Economics, 24,* 131-142.

Schacter, J.P., Franklin, R.S., & Perry, M.J. (2003). *Migration and Geographic Mobility in Metropolitan and Nonmetropolitan America: 1995 to 2000.* (U.S. Census Bureau Census 2000 Special Reports CENSR-9). Washington, D.C.: Government Printing Office.

Schacter, J.P., Jensen, L., & Cornwell, G.T. (1998). Migration, residential mobility, and poverty in rural Pennsylvania. *Rural Development Perspectives, 13,* 40-45.

Schwartz, A. (1973). Interpreting the effect of distance on migration. *The Journal of Political Economy, 81,* 1153-1169.

Schwartz, A. (1976). Migration, age, and education. *The Journal of Political Economy, 84,* 701-720.

Semple, R.K. & Gollege, R.G. (1970). An analysis of entropy changes in settlement patterns over time. *Economic Geography, 46,* 157-160.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal, 27,* 379-423, 623-656.

Shryock, H.S. & Siegel, J.S. (1976). *The methods and materials of demography*. San Diego: Academic Press.

Siegel, P.B, Johnson, T.G., & Alwang, J. (1995) Regional economic diversity and diversification. *Growth and Change, 26,* 261-284.

Sjaastad, L.A. (1962). The costs and returns of human migration. *The Journal of Political Economy, 70,* 80-93.

Snyder, D. & Kick, E.L. (1979). Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions. *American Journal of Sociology, 84,* 1096-1126.

Straathof, S.M. (1997). Shannon's entropy as an index of product variety. *Economics Letters, 94,* 297-303.

Strauss, J.M. (1944) The entrepreneur: The firm. *The Journal of Political Economy, 52,* 112-127.

Taylor, M.P. (1996). Earnings, independence or unemployment; Why become self employed? *Oxford Bulletin of Economics and Statistics, 58*, 253–265.

Thurik, A.R., Carree, M.A., van Stel, A. & Audretsch, D.B. (2008). Does self-employment reduce unemployment? *Journal of Business Venturing, 23*, 673-686.

Tutzauer, F. (2007). Entropy as a measure of centrality in networks characterized by path-transfer flow. *Social Networks, 29*, 249–265.

U.S. Bureau of Economic Analysis. (2010, April). *Local area methodology*. Retrieved November 22, 2010, from http://www.bea.gov/regional/pdf/lapi2008/employment.pdf

U.S. Bureau of Economic Analysis (BEA) Regional Economic Accounts (REIS). (2010). *Regional economic accounts.* Retrieved October 20, 2010, from http://www.bea.gov/regional/

U.S. Census Bureau. (2000a). *County Business Patterns: 2000.* Retrieved September 25, 2010, from http://www.census.gov/econ/cbp/download/00_data/index.htm

U.S. Census Bureau. (2000b). *County-to-county migration flow files.* Retrieved September 25, 2010, from http://www.census.gov/population/www/cen2000/ctytoctyflow/index.html

U.S. Census Bureau. (2000c). *Summary files 1 and 3.* Retrieved Sep 14, 2010, from http://dataferrett.census.gov/

U.S. Census Bureau. (2001). *Centers of population computation for 1950, 1960, 1970, 1980, 1990, and 2000.* Retrieved Jan 27, 2011, from http://www.census.gov/geo/www/cenpop/calculate2k.pdf

U.S. Census Bureau, Current Population Survey (CPS). (2010). *Geographic Mobility/Migration.* Retrieved Sep 19, 2010, from http://www.census.gov/population/www/socdemo/migrate.html

U.S. Census Bureau, Geography Division. (2010). Mean center of the population for the United States: 1790 to 2010. Retrieved April 4, 2011, from http://www.census.gov/geo/www/2010census/centerpop2010/centerpop_mean2010.pdf

U.S. Department of Agriculture (USDA) Economic Research Service (ERS). (2004). *Measuring rurality: Rural-urban continuum codes.* Retrieved Sep 23, 2010, from http://www.ers.usda.gov/briefing/rurality/ruralurbcon/

Venables, A.J. (2006). Shifts in economic geography and their causes. *Economic Review, Federal Reserve Bank of Kansas City, 4th Quarter,* 61-83.

Wasserman, S. & Faust, K. (1994). *Social network analysis.* New York: Cambridge University Press.

Wellman, B. (1983). Network analysis: Some basic principles. *Sociological Theory, 1,* 155-200.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica, 48,* 817–838.

Wolpert, J. (1965). Behavioral aspects of the decision to migrate. *Papers in Regional Science, 15,* 159-169.