Parajuli, Jitendra; Haynes, Kingsley

Conference Paper
An Exploratory Analysis of New Firm Formation in New England

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Parajuli, Jitendra; Haynes, Kingsley (2013) : An Exploratory Analysis of New Firm Formation in New England, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
http://hdl.handle.net/10419/123894

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
An Exploratory Analysis of New Firm Formation in New England
Jitendra Parajuli
Kingsley E. Haynes*

*Corresponding author
School of Public Policy, George Mason University
3351 Fairfax Drive, MS 3B1, Arlington, VA 22201
khaynes@gmu.edu
1-703-993-2280 (O)
1-703-993-2284 (F)

Abstract
This study examined the distribution of new firm formation in New England from 1999 through 2009. Using discrete entropy and entropy decomposition, it was found that single-unit firm births are spatially dispersed. The distributional patterns do not vary substantially across the study period. However, some industry sectors, such as information, and finance and insurance, were more concentrated than sectors, such as health care and social assistance, and accommodation and food services. Although the extent of new firm formation is generally associated with business cycles, economies of scale associated with the regions are more likely to explain the distributional patterns of new firm formation.
Keywords: new firm formation, entropy, distribution patterns, spatio-temporal patterns, entrepreneurship, financial crisis, recession, New England

JEL classification: R30, C40

Introduction

In 2007, the US experienced an economic crisis that is often compared with the Great Recession. High risk lending, regulatory failure, inflated credit ratings, and investment bank abuses were considered to be the drivers of the recession (US Senate, 2011). The US government took various initiatives to minimize as well as mitigate its effect (see Simkovic, 2009).

Entrepreneurs play a crucial role in economic growth (Schumpeter, 1934, 1942), and through new firms they create new jobs and distribute wealth (Kirchhoff, 1994). Although the US has the right mix of factors for entrepreneurial activities (Acs & Armington, 2006; World Economic Forum, 2012; Fairlie, 2013), the recession of 2007 had a detrimental impact on US entrepreneurship—the number of entrepreneurial ventures, especially new firm formation, declined and self-employment decreased (see Shane, 2011).

Various studies have examined the causes and impacts of the recession of 2007-08, but the distributional patterns of new firm formation have not been examined. This paper is an attempt to understand the distribution of new firm formation in New England from 1999 through 2009. The results are likely to show the variability of distribution of new firms and that regional planners and policy makers are likely to benefit from the results of this study.

This paper is arranged as following. Next section presents theoretical background and research questions of this study followed by the methodological framework and data. Thereafter, the results are presented. Finally, the concluding section summarizes the findings and provides policy implications and the future direction of research.
Theoretical background and research questions

Recession is the contraction of business cycle and is associated with the fall of real gross domestic product and employment. Since there are various types of business cycles—Juglar, Kitchin, Kuznets, and Kondratiev—and fundamental differences exist among macroeconomic performance and business cycles, it is a challenge to identify the short-run and long-run economic behavior and predict the factors that cause economic fluctuations (see Abarmovitz, 1961; Tvede, 2006).

Economic crises with dire consequences were not uncommon in the history (see Kindelberger & Aliber, 2011), but a unifying theory that is helpful in understanding the causes and remedial measures is missing. Keynes ([1936] 1957) posited the rise and fall of aggregate demand as the source of economic fluctuations. Although the supply and demand could balance out at equilibrium, it would not deliver full employment. Thus, government needs to deliberately finance on public works to create jobs. On the contrary, Ludwig von Mises suggested that the intervening banking policy—underbidding the interest rates and expanding the credit—is the source of business cycle. Policies that establish interest rates, wage rates, and commodity prices as opposed to the market mechanism not only invite destruction and chaos, but also hamper economic progress (see von Mises, 2006). According to Schumpeter (1934, 1939), an economy is a self-perpetuating system that experiences booms and busts. This cyclical phenomenon is mainly caused by innovations and is non-monetary. Other theoretical frameworks proposed by Friedman (1970), Lucas (1975), and Kydland and Prescott (1982) are also adopted in understanding economic fluctuations and their implications.

Minsky (1976) proposed the theory of financial crises which suggests that hedge finance, speculative finance, and Ponzi schemes are vulnerable to the market condition and become liabilities to firms. Because of their vulnerabilities, any shock in the economy induces financial disturbances and thus affects it. In addition, there are also other factors, such as coordination failures, moral hazards, regulatory mechanism, and liquidity crunch, that contribute to financial crises (see Brunnermeier, 2009; Crotty, 2009; Goldstein & Razin, 2013).
After the 2007 financial crisis, the US government created a bipartisan committee, the *Financial Crisis Inquiry Committee*, to study its causes. According to the Committee, the crisis was a result of human action and inaction that was tied to issues, such as financial regulation and markets; failures of corporate governance and risk management at financial institutions; a combination of excessive borrowing, risky investments, and lack of transparency; and a systemic breakdown in accountability and ethics (Financial Crisis Inquiry Commission, 2011). Similarly, the *Levin-Coburn Report* of the US Senate attributed high risk lending, regulatory failure, inflated credit ratings, and investment bank abuses as the major drivers of the financial crisis (US Senate, 2011). Others argued that that both the financial sector and the government were the enablers of the turmoil (see Krugman, 2009; Reinhart & Rogoff, 2009; Kindelberger & Aliber, 2011).

Entrepreneurs play a crucial role in economic growth (Schumpeter, 1934, 1942), and through new firms they create new jobs and distribute wealth (Kirchhoff, 1994). Empirical studies in the field of entrepreneurship research have suggested that population growth (Guesnier, 1994; Audretsch & Fritsch, 1994; Reynolds, Miller, & Maki, 1995); income growth (Armington & Acs, 2002; Lee, Florida, & Acs, 2004); human capital (Armington & Acs, 2002); financial capital (Sutaria & Hicks, 2004); social capital, and cultural diversity (Saxenian, 2002; Lee et al., 2004; Audretsch, Dohse, & Niebuhr, 2010; Hart & Acs, 2011), in general, are positively associated with new firm births. However, the relationships between establishment size and new firms (Armington & Acs, 2002; Fritsch & Falck, 2002; Sutaria & Hicks, 2004), and employment size and new firms are not consistent (Storey, 1991; Reynolds et al., 1995; Fritsch & Falck, 2002; Sutaria & Hicks, 2004).

A well-functioning financial system is important for economic growth and development (see Levine, 1997; Aghion, Howitt, & Mayer-Foulkes, 2005). According to Schumpeter (1934), since innovations are important for the growth of a capitalist society, banking system is crucial for identifying and funding promising innovations. In today’s economic environment, capital sources extend beyond the conventional banking system and that entrepreneurs can obtain financial capital sources, such as venture capitalists, personal savings, and social networks. However, Rajan and Zingales (1998) suggested that while financial development can enhance
innovation and growth indirectly, the imperfection of financial market can have detrimental impact on investment and growth.

A number of regional development theories on spatial distribution of economic activities has been proposed. The theory of location was conceptualized by Johann-Heinrich von Thunen according to which agricultural activities locate in the proximity of transportsations facilities and the locational factor is also related to the theory of rent and land-use. Isard (1949), on the other hand, pointed out that the general equilibrium analysis fails to account for spatial dimension and neglects transportation and spatial costs on the distribution of economic activities in space and should consider the spatial array of economic activities with the geographic distributions of inputs and outputs and prices and costs. More recent works of Krugman (1991, 1998), Markusen (1996), Ellison and Glaeser (1997), and Fujita, Krugman, and Venables (1999) also explain the configurations and regional variations of economic activities.

According to Arthur (1990), new firms that spin-off from existing firms are found in close geographic proximity to their genetic parent. Porter’s theory of competitive advantage that emphasizes the important of geographic clusters suggested that new businesses locate themselves in clusters that supply specialized inputs and required specialized infrastructure (Porter, 2000). In addition, Arbia (2001) noted that there is a higher chance of a new firm to locate in the neighborhood of existing firms. Thus, whether in West-Germany (Bade & Nerlinger 2000) or in the US (Renski, 2009) new firms locate themselves in regions that are beneficial for growth and survival.

The purpose of this paper is to understand the distribution patterns of new firm formation in New England—Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont—from 1999-2009. Based on the theoretical background and empirical evidence, namely economic fluctuations and the financial crisis of 2007, the study proposes to find answers to the following questions:

1. How are new firm (single-unit) births distributed in New England from 1999-2009?
2. Are the distributions of new firms comparable across industry sectors?

3. Are the distributions of new firms comparable to the distributions of other economic indicators?

**Method and data**

The Gini coefficient, the Herfindahl index, the ogive index, and entropy are some of the commonly used techniques for measuring economic diversification. The Gini coefficient is used for characterizing inequalities. It is defined as a ratio of the area under the Lorenz curve, and its value ranges from zero (complete equality) to one (complete inequality). Although the Gini coefficient is a popular measure of income and wealth concentrations, it is not additive decomposable. That is, it is not possible to derive within and between values (see Theil, 1967; Shorrocks, 1980).

The Herfindahl index, also known as the Herfindahl-Hirschman index, is another measure of concentration generally used for measuring industrial concentration (see Hirschman, 1964). It is defined as the sum of the squares of all values of the variable of interest. The maximum value of the Herfindahl index is one (complete concentration) and the minimum is the inverse of all the observations. Unlike the Gini coefficient, the Herfindahl index is (multiplicative) decomposable (see Theil, 1967). Since the measure gives heavier weight to the higher values of the variable, this index can often misrepresent the actual concentration.

The ogive index is a measure of variance of a variable against a uniform distribution. Often, employment is used as the benchmark. However, with the assumption of uniform employment share across the industry sectors and changes arising from the shift in national economy, it would be difficult to defend the use of ogive as a measure of diversity index (see Bahl, Firestine, & Phares, 1971; Wasylenko & Erickson, 1978).

Entropy is a measure of uncertainty and the Shannon’s discrete entropy is defined as (Shannon, 1948):

\[
H = \sum_{i=1}^{n} p_i \log \left( \frac{1}{p_i} \right) = - \sum_{i=1}^{n} p_i \log p_i
\] (1)
where $n$ is the number of discrete observations and $\{p_i\}$ are discrete probabilities. $H$ is a continuous and an increasing function of $n$. Its value ranges from 0 to $\log n$ ($H \sim H_{\text{max}}$). As $H$ tends to zero, spatial concentration goes toward maximum, and as $H$ tends to $\log n$, spatial concentration goes toward minimum.

Relative entropy is a useful measure for comparing distributions and assumes values between 0 and 1. It is defined as:

\[
\text{Relative } H = \frac{H}{\log n}
\]  

(2)

Entropy can be decomposed into two components—(1) between-set and (2) within-set entropies (Theil, 1967)—and the decomposed entropy can be written as:

\[
H = \sum_j p_j \log \frac{1}{p_j} + \sum_j p_j \left[ \sum_{i \in j} \left( \frac{p_i^{(j)}}{p_j} \right) \log \left( \frac{1}{p_i^{(j)}/p_j} \right) \right] = H_j + \sum_j p_j H_i^{(j)}
\]

(3)

where $H_j$ is the between-set entropy and $\sum_j p_j H_i^{(j)}$ is the average within-set entropy for smaller observations $i$ embedded in larger observations $j$. In addition, Equation (1) and Equation (3) must be equal.

The between-set relative entropy is then calculated as:

\[
\text{Relative } H_j = \frac{H_j}{\log j}
\]

(4)

Likewise, the average within-set relative entropy is:

\[
\text{Relative } H_{n/j} = \frac{H_{n/j}}{\log(n/j)}
\]

(5)

Information gain is a measure that can be used for comparing two distributions and the expected information gain is written as:

\[
I = \sum_{i=1}^{n} q_i \log \left( \frac{q_i}{p_i} \right)
\]

where $\{p_i\}$ is the set of prior probabilities and $\{q_i\}$ is the set of posterior probabilities. If the distributions are similar, $I$ tends toward zero.
Entropy measure also has a number of limitations—ambiguous definitions, the assumption of uniform distribution, and spatial boundary definition (see Theil, 1967; Batty, 1974; Haynes, Phillips, & Mohrfeld, 1980). However, since entropy is asymptotically normally distributed, it is commonly used for measuring diversity (see Haynes & Storbeck, 1978; Pannell, 1988; Wheeler, 1990; Kulkarni, Stough, & Haynes, 1999; Feser, Hewings, & Mix, 2013).

According to Acs and Armington (2006) the spirit of entrepreneurship is reflected by the creation of new firm instead of expansion of an existing firm. Along this line, entropies of new firm formation will be measured for New England from 1999-2009. The data for single-unit new firm births are from the US Census Bureau. Since the relationship between unemployment and entrepreneurial venture is indeterminate, unemployment will be used for measuring the expected information gain. The unemployment data are from the US Bureau of Labor Statistics.

**Results**

Figure 1 shows the standard deviation of the distribution of total single-unit firm births in New England in 2002.
In 2002, although Cumberland County in Maine and Fairfield, New Haven, and Hartford counties in Connecticut experienced a more than average firm births in New England, Massachusetts had the largest number of single-unit firm births. During the study period, this trend was generally true.

![Figure 2: Total single-unit firm births in New England (1999-2009)](image)

Figure 2 shows the total number of new firm formation (in logarithm) from 1999-2009. It should be noted that after 2005, and notably from 2007 onwards, the number of new firm formation slowly started to decline.

<table>
<thead>
<tr>
<th>Year</th>
<th>Shannon’s entropy</th>
<th>Relative entropy</th>
<th>Between-region entropy</th>
<th>Within-region entropy</th>
<th>Between-region relative entropy</th>
<th>Average within-region relative entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>3.576765</td>
<td>0.850660</td>
<td>1.519784</td>
<td>2.056981</td>
<td>0.848208</td>
<td>0.852482</td>
</tr>
<tr>
<td>2000</td>
<td>3.568449</td>
<td>0.848683</td>
<td>1.512870</td>
<td>2.055579</td>
<td>0.844349</td>
<td>0.851901</td>
</tr>
<tr>
<td>2001</td>
<td>3.602396</td>
<td>0.856756</td>
<td>1.540509</td>
<td>2.061887</td>
<td>0.859774</td>
<td>0.854515</td>
</tr>
<tr>
<td>2002</td>
<td>3.605551</td>
<td>0.857506</td>
<td>1.535117</td>
<td>2.070434</td>
<td>0.856765</td>
<td>0.858057</td>
</tr>
<tr>
<td>2003</td>
<td>3.620132</td>
<td>0.860974</td>
<td>1.555141</td>
<td>2.064991</td>
<td>0.867941</td>
<td>0.855801</td>
</tr>
<tr>
<td>2004</td>
<td>3.618029</td>
<td>0.860474</td>
<td>1.552844</td>
<td>2.065185</td>
<td>0.866659</td>
<td>0.855882</td>
</tr>
</tbody>
</table>
Table 1: Entropy measures for total single-unit firm births in New England (1999-2009)

Table 1 shows various entropy measures across years. While there was a declining trend in the number of new firms after 2005, all the measures suggest that the distribution patterns of new firm formation were almost consistent in New England. In general, firm formation was dispersed. The between-region entropies suggest that the distribution patterns of new firm births were more concentrated across states. However, the average within-region entropies indicate that, on average, firm formation patterns dispersed more within states than across states. Thus, although the number of new firms was declining, there was a minimal change in the actual distribution pattern of firm births (see also Figure 3).

![Graph](image-url)

Figure 3: Relative entropies for total single-unit firm births in New England (1999-2009)

However, it should also be noted that while there is no dramatic change in the relative entropy values, their decline after 2007 suggests that firm births were more concentrated in New England.
The total number of single-unit firm births was disaggregated into manufacturing and service sectors. The service sector included—(1) Wholesale trade; (2) Retail trade; (3) Transportation and warehousing; (4) Information; (5) Finance and insurance; (6) Real estate and rental and leasing; (7) Professional, scientific, and technical services; (8) Management of companies and enterprises; (9) Administrative and support and waste management and remediation services; (10) Educational services; (11) Health care and social assistance; (12) Arts, entertainment, and recreation; (13) Accommodation and food services; (14) Other services (except public administration).

The general trend in the manufacturing and sector sectors was that new firm births were declining, especially after 2007. However, various entropy measures for both these sectors
(Table 2) indicate that the distribution patterns of new firm formation were not much different from the distribution patterns of total new firm births.

Figure 4: Relative entropies for single-unit firm births (Manufacturing)

Figure 5: Relative entropies for single-unit firm births (Service)

The temporal changes of relative entropies for the manufacturing and service sectors are shown in Figure 4 and Figure 5. Note that in the aftermath of the financial crisis of 2007, the distribution patterns of new firm formation became more concentrated in New England.
### Table 3: Entropy measures for single-unit firm births (Information, and finance and insurance sectors)

The service sector data was further disaggregated to individual service sectors. Table 3 shows entropies of the information, and finance and information sectors. Firm births in both these sectors were less dispersed than at the aggregate or the manufacturing and service sector levels.

![Figure 6: Relative entropies for single-unit firm births (Information)](image)
All the relative entropies in both the sectors dropped after 2007. However, the continuous decline in the entropy values of the finance and insurance sector, especially since 2006, suggests that firm formation in this sector was more localized than firm formation in other sectors.

Figure 7: Relative entropies for single-unit firm births (Finance and insurance)

Figure 8: Unemployed labor force
Within the service sector, the professional, scientific, and technical services sector also showed similar firm births distribution patterns. However, single-unit new firm formation in the real estate and rental and leasing, health care and social assistance, and accommodation and food services sectors were more dispersed.

<table>
<thead>
<tr>
<th>Year</th>
<th>Unemployment (No lag)</th>
<th>Unemployment (One year lag)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>0.055696</td>
<td>0.061255</td>
</tr>
<tr>
<td>2000</td>
<td>0.046213</td>
<td>0.057738</td>
</tr>
<tr>
<td>2001</td>
<td>0.042695</td>
<td>0.045875</td>
</tr>
<tr>
<td>2002</td>
<td>0.047148</td>
<td>0.038611</td>
</tr>
<tr>
<td>2003</td>
<td>0.053654</td>
<td>0.050270</td>
</tr>
<tr>
<td>2004</td>
<td>0.052396</td>
<td>0.051721</td>
</tr>
<tr>
<td>2005</td>
<td>0.048540</td>
<td>0.053066</td>
</tr>
<tr>
<td>2006</td>
<td>0.048646</td>
<td>0.049266</td>
</tr>
<tr>
<td>2007</td>
<td>0.045828</td>
<td>0.047929</td>
</tr>
<tr>
<td>2008</td>
<td>0.048304</td>
<td>0.043554</td>
</tr>
<tr>
<td>2009</td>
<td>0.041410</td>
<td>0.048105</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations

Table 4: Expected information gain

Figure 8 shows the unemployment trends in New England. The expected information gain was calculated to compare the distribution patterns of firm formation and unemployment. Unemployment was also lagged by a year to examine the differences in the distribution patterns of firm births. It was found that whether the unemployment was lagged or not, very small values of the expected information gain suggest that firm formation and unemployment patterns more or less follow each other.
Conclusion

The purpose of this paper was to examine the distribution patterns of new firm formation in New England. It was found that single-unit firm births are generally dispersed across New England between 1999 and 2009. However, distributions of firm births in sectors, such as information, and finance and insurance, are more concentrated than sectors, such as health care and social assistance, and accommodation and food services. Moreover, within-region entropies suggest that firm births dynamics could be localized in some counties embedded in their respective states. The values of expected information gains suggest the existence of scale economy and business cycle. In other words, regions that have a higher unemployment experience a higher number of firm births and that phenomenon can be attributed to economic fluctuations. Thus, government policy to foster economic competitiveness in a region through new firms should take into account the spatio-temporal distributional dynamics of new firm formation, and scale economy and business cycle.

Future research should include spatial entropy extension. In the case of data availability, other economic regions should be included and regional comparison should be made. In addition, econometric analysis of entropy could also be done.

References

An Exploratory Analysis of New Firm Formation in New England
Parajuli and Haynes


