12-1-2005

Regional Economic Issues: Determinants of Recession Duration and Determinants of New Firm Formation

Angela M. Kuhlmann
University of Nebraska at Omaha

REGIONAL ECONOMIC ISSUES: DETERMINANTS OF RECESSION DURATION AND DETERMINANTS OF NEW FIRM FORMATION

A Thesis

Presented to the

Department of Economics

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts in Economics

University of Nebraska at Omaha

By

Angela M. Kuhlmann

December 2005

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

THESIS ACCEPTANCE

Acceptance for the faculty of the Graduate College, University of Nebraska, in partial fulfillment of the requirements for the degree of Master of Arts in Economics, University of Nebraska at Omaha.

Committee

Thomas R. Gulledge - Geography

Mark E. Wohar

Chairperson

Date 11-29-05

REGIONAL ECONOMIC ISSUES: DETERMINANTS OF RECESSION DURATION AND DETERMINANTS OF NEW FIRM FORMATION

Angela M. Kuhlmann, MA

University of Nebraska at Omaha, 2005

Advisor: Christopher Decker

ABSTRACT

Why are some regional economies able to outperform those of other regions? Using state-level data in two separate analyses, this thesis shows that industrial diversity is one potential answer. Industrial diversity is calculated using the Herfindahl index. In the first model, duration analysis on state recessions occurring between 1979 and 1996 indicates that an increase in industrial diversity is associated with shorter recessions. Other determinants of recession duration include unemployment, change in real income per capita, proportion of non-white workers, total population, and change in population growth. In the second model, regression analyses on 2001 firm formation rates show that higher levels of industrial diversity are associated with higher rates of new small firm formation. Other determinants of small firm formation include education, availability of financing, average size of existing establishments, and presence of environmental hazardous waste sites. Based on these results for industrial diversity, state policies aimed at increasing diversity appear to be justifiable.

ACKNOWLEDGEMENTS

This thesis is the product of many hours of both discussions with Professor Decker and independent readings and research. It would not have been possible without the guidance, assistance, and feedback from Professor Decker as well as Professor Wohar and Professor Gildersleeve. I am forever grateful for their time and efforts in advising me through this research project and want to take this opportunity to thank them. Thank you also to my husband and parents for their continued support and encouragement in helping me attain my goals.

TABLE OF CONTENTS

FIGURES AND TABLES

Page:

List of Figures

List of Tables

List of Appendix Tables

I. Introduction

In recent years, many states have been dealing with budget dilemmas, and cuts at the national level have further worsened their financial situation. This has raised many questions and increased the focus on economic performance at the sub-national level. Aggregate national data often obscure the economic reality in regional and local environments due to significant regional variation in economic conditions, cycle dynamics, and reactions to monetary policy (Wall and Zoega 2004). While the considerable variation among state economies is evident, the underlying reasons for the variability are less clear. Why are some state economies able to consistently and significantly outperform those of other states? What factors are most important in determining a state's level of economic performance?

Economic performance can be evaluated in various ways using a wide range of macroeconomic measures. Some of the most common measures include unemployment, per capita income, labor productivity, and job creation. This research effort focuses on two less commonly used, but no less revealing, measures of economic performance: duration of recessions and new firm formation. While the two ideas seem mostly unrelated, previous literature has made some intriguing ties between recessions and new business start-ups.

For example, there is a debate over whether new business start-ups are motivated by the "push" of recessions or whether they are more influenced by the "pull" of expansions (Armstrong and Taylor 2000, p. 270). Does the increased unemployment of

recessions push more potential entrepreneurs into starting their own businesses as an alternative to being unemployed? Or does the increased income of expansions pull more entrepreneurs away from their employers to exploit market opportunities for themselves?

Examining the relationship in reverse, Audretsch and Keilbach (2004) use entrepreneurial capital, as measured by new business creation, as an independent variable in explaining state variation in output and they find it to be positive and statistically significant. Extending their logic, it is possible that a state with more entrepreneurial capital would be more creative and less risk-averse, therefore more adaptive to recessions and able to rebound from economic downturns more quickly.

One independent variable of particular interest for explaining either the duration of recessions or the number of small business startups is industrial diversity.[1] This issue of diversity as related to business cycles is interesting because there are conflicting theories surrounding it. One argument, based on traditional Ricardian trade theory, is that in order to grow, a region should not diversify but rather specialize in whichever goods it has a comparative advantage in producing (i.e. those goods that are relatively cheaper to produce). That region should then trade with other regions for goods that are relatively more costly for them to produce themselves, and thus they will be better off. However, the counter argument is that a region should diversify in order to increase the stability of its economy. Regions experience different shocks to output depending on their industrial mix because economic shocks tend to affect certain industries more deeply. It is

[1] When explaining the length of recessions, this study measures industrial diversity using the Herfindahl Index, defined as $\Sigma(E_{is}/E_s)^2$ where E_{is}/E_s is the employment share of industry *i* in state *s*. The other common measure of industrial diversity is the entropy index, defined as $\Sigma(E_{is}/E_s)\ln(E_s/E_{is})$.

hypothesized that more industrially diverse states experience shorter recessions because the greater number of industries is able to absorb the cyclical unemployment more quickly.

Could industrial diversity also influence entrepreneurs and new firm formation? If a state is more industrially diverse, there may be a greater potential for new firms to compete and to exploit niche markets due to the interaction among the numerous industries. However, there are counter arguments here as well. In states that are more industrially specialized, a few large growth industries may actually offer more opportunities for new start-ups due to spillover effects, as we have seen in the computer industry. While this relationship between new firm formation and industrial diversity has not been extensively researched in the current literature, it is explored in the second section of this paper. First, the issue of recession duration and industrial diversity is examined.

II. Determinants of State-Level Recession Duration

1. Background

According to the National Bureau of Economic Research (NBER), a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally measured by declines in real GDP, real income, employment, industrial production, and wholesale-retail sales. Typically, a downturn is considered a recession if real GDP declines for at least two consecutive quarters (NBER 2005). A recession is a broad measure of economic performance, and therefore this study examines the duration of recessions rather than the duration of unemployment alone.

There are several reasons why examining the duration of recessions is interesting and informative. A recent trend shows that states are increasingly relying on their "rainy day funds" to meet budget shortfalls in the short run in order to avoid raising taxes and/or cutting expenditures (Wagner and Elder 2004). The longer the duration of a recession, the more likely it is that a state's rainy day funds will dry up. Moreover, the longer a recession continues, the longer it will be before a state has extra revenue to begin replenishing their rainy day funds. If revenue fails to rebound quickly enough, spending cuts and/or tax increases may have to be larger than previously necessary in order to replenish these emergency funds.

From a labor perspective, longer recessions lead to longer durations of unemployment. The duration of unemployment more accurately reflects the welfare of workers rather than a measure of whether they are unemployed or not at some given point

in time. As the duration of unemployment lengthens, more workers become discouraged and drop out of the labor force, leading to corresponding declines in consumption that further exacerbate the impacts of a recession.

The longer a state recession persists, the more likely it is that workers may begin to migrate to more prosperous states. Research has shown that gross migration flows tend to fall during national recessions due to factors such as increased difficulties finding a new job or lower earnings in previously attractive destinations (Armstrong and Taylor 2000, p. 158). However, net migration flows among regions are more complex. Cebula (2005) finds that both per capita income and expected per capita income have statistically significant positive impacts on interstate migration flows. If a recession is very short, it is unlikely that the expected earnings differential would outweigh the immediate costs of migrating. On the other hand, if relatively large disparities between states persist, then a longer recession will eventually induce workers to migrate as the gap between job opportunities and earnings increases between their home state and a more prosperous destination state.

Along similar lines, if a state recession persists, new firms may be less likely to locate in that state where the business climate appears to be depressed. While this has not been formally researched, it is theoretically a possibility. Firms may recognize that certain states recover from economic downturns more quickly and this could affect their location decisions. Not only do some states recover more quickly, but they may not experience the initial downturn as soon as other states with less diversity in their economic structure.

Many of these reasons for studying recession duration rather than recession depth are also good reasons for studying recessions in general rather than expansions. Arguably, prolonging economic expansions requires less drastic efforts than turning the business cycle around to end recessions. Policymakers can take measures to keep an expansion rolling, but more aggressive policies may be necessary to change the economic trends that are causing a recession and turn the business cycle around. Policymakers are particularly interested in ending recessions for the reasons mentioned above such as reducing the number of discouraged workers dropping out of the labor force and limiting outward migration that may result from prolonged recessions.

The goals of this paper are (1) to test whether state-level recessions are more likely or less likely to persist as they increase in duration, and (2) to investigate what factors and characteristics of state economies influence the duration of recessions, testing the influence of industrial diversity in particular. Regarding the first goal, the main finding is that most recessions are more likely to end as their duration increases given that they have not ended already. Research for the second goal revealed that numerous factors are significantly related to longer recessions. The factors associated with longer recessions include decreases in industrial diversity, high unemployment, decreases in real income per capita, high proportions of non-white workers, small total populations, and increases in population growth.

The rest of this chapter is organized as follows. Section 2 discusses the literature related to state economic performance and industrial diversity. The data and econometric methodology for duration analysis are explained in Section 3, addressing both non-

parametric and full parametric analysis. The full parametric estimation results are presented and analyzed in Section 4 followed by general conclusions and policy implications in Section 5.

2. Literature Review

Based on numerous literature searches, no paper appears to have been written with the exact objective of this chapter, that is to study recession duration with a particular focus on the influence of industrial diversity. However, various studies have been performed investigating the effects of state-level industrial diversity on other measures of economic performance. Their approaches and results were helpful in defining the full parametric model utilized later.

Nearly all studies examining regional variations in economic performance find industrial diversity to have a statistically significant effect. The difference lies in how they measure and represent industrial diversity in their analyses. Many studies use industry proportions to indicate the predominance of certain industries. For example, Garcia-Mila and McGuire (1993) use the shares of employment in certain industries, like farming and non-durable manufacturing, as determinants in explaining employment growth rates and variability at the state level. After controlling for variation in industry growth at the national level and for the composition of fast and slow growth industries at the state level, they find that industrial mix still has a significant effect on state growth rates over the period 1969-1985.

Owyang et al. (2004) also use employment shares of certain industries as explanatory variables in their analysis of growth rates. They examine growth rates within recessions and expansions separately using monthly state coincident data from Crone (2002). The general conclusion from Owyang et al. is that industrial mix affects growth rates during recessions but not during expansions. In fact, their results indicate that differences in recession growth rates are predominately influenced by industrial mix, while expansion growth rates are related to differences in demographics and not to industrial composition.

Instead of employment shares, Carlino and Sill (2001) use the share of total output accounted for by certain industries as an independent variable in their regressions. They analyze cycle and trend growth rates separately, and they find that industry mix has a differential effect on real income growth. However, the implications of their study are not clear because the effect of industry mix is sometimes positive and sometimes negative, depending on which region and which growth rate (cycle or trend) they examine.

There are many alternatives to using industry shares to measure diversity, as outlined in Wundt (1992). For example, a "percent durables index" is often used to reflect the fact that durable goods have a higher income-demand elasticity and therefore can predict cyclical instability. The national average index is sometimes used to capture the deviations in state industry shares from the national shares. As mentioned in footnote number 1 in the introduction, the entropy index is a very common measure of industrial diversity that equals 0 if a state has perfect concentration or equals 1 if a state has perfect

diversity. Wundt also discusses portfolio variance as a newer tool to measure industrial diversity that is based on detrended industry employment shares as well as interindustry employment variances. Wundt compares the explanatory power of all of these various measures in predicting regional cyclical employment behavior. He finds most of the measures to be statistically significant, and all of them indicate that greater industrial specialization is associated with greater instability.

Finally, the Herfindahl index is a very commonly used measure of industrial diversity. Simon and Nardinelli (1992) and Izraeli and Murphy (2003) both use this approach. Simon and Nardinelli find that in all years studied, except during the Depression years of 1930 and 1931, more industrially diversified cities experienced lower unemployment. They attribute this to the portfolio effect, or in other words, workers can find employment more easily when there is a greater number of industries in the region.[2] During 1930 and 1931, more diversified cities actually experienced higher unemployment. The authors speculate that wages adjust downward more quickly in specialized regions, therefore preventing lay-offs, because of the limited probability that workers will quit and change industries. Only during these two years though, did the wage adjustment effect outweigh the portfolio effect.

Similarly, Izraeli and Murphy (2003) find that a lower Herfindahl index (indicating higher industrial diversification) is associated with lower unemployment, supporting their thesis that diversification can reduce unemployment. They rely on state-

[2] This portfolio effect is basically the same idea as agglomeration economies. A large number of geographically concentrated economic activities, usually spread across multiple industries, allows labor to be reallocated more efficiently when one industry experiences a downturn.

level data and according to their results, a 10-point decrease in the Herfindahl Index is associated with almost a tenth of a percentage point decrease in the unemployment rate. Therefore, the authors support state policies targeting a more diversified industrial base as a means to provide more employment security during economic downturns.

Every measure of industrial diversity used in the literature and reviewed above has its advantages and disadvantages. The measure chosen for this paper is the Herfindahl index, which is most similar to the entropy index in terms of how it is calculated.[3] Because the Herfindahl index is a more comprehensive measure of the diversity issue of interest, this research relies on the Herfindahl index as Simon and Nardinelli (1992) and Izraeli and Murphy (2003) do. Furthermore, this study most closely follows Izraeli and Murphy in terms of the explanatory variables used in the full parametric estimation section. However, instead of examining unemployment, the dependent variable here is the duration of recessions.

3. Data and Econometric Methodology

The state recession data used in this analysis come from Crone (2002). Crone calculated his state coincident indexes based on Stock and Watson (1989).[4] According to Crone, the advantage of the Stock and Watson method is that it provides a single measure of a state's economy by combining several monthly indicators. Crone specifically bases

[3] Another similar alternative not mentioned previously would be to measure industrial diversity with a location quotient Herfindahl index. In other words, the variable could be calculated as the state Herfindahl index relative to the national Herfindahl index.

[4] Stock and Watson's index is the latent factor estimated in a dynamic single-factor model using the Kalman filter.

his measure on four indicators: (1) nonagricultural employment, (2) the unemployment rate, (3) average hours worked in manufacturing, and (4) real wage and salary disbursements. These coincident indexes are particularly useful for comparing the length, depth, and timing of state recessions because they utilize consistent economic measures across all 50 states and they are available at a greater frequency than gross state product, for example, which is only available on an annual basis. The earliest time period covered by Crone is 1979:Q1 so this analysis uses all recessions from that quarter through 1997:Q4, at which point the industrial classification system changed significantly enough to affect analyses spanning the time periods of the two different systems.

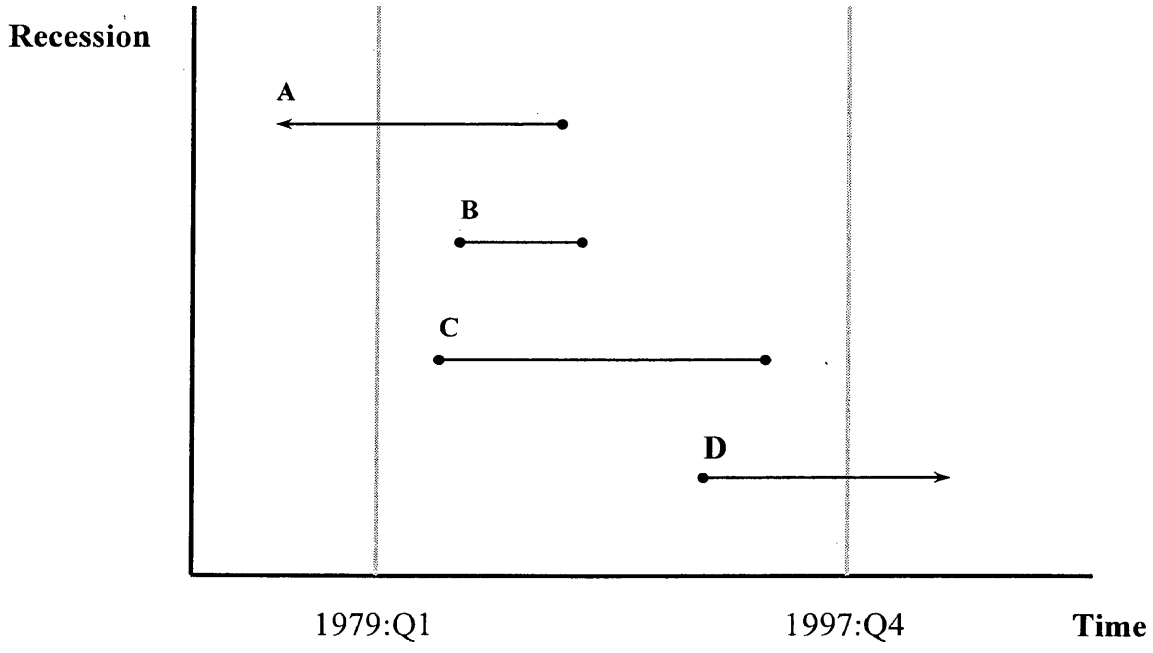
Table 2.1 lists some descriptive statistics for these recession data. Two recessions with durations of 30 and 52 quarters (for Hawaii and Alaska respectively) seemed to be outliers. However, these two recessions did not end during the time period examined. Therefore, instead of removing them from the sample as would typically be done with outliers, they are treated as censored observations, which is explained below.

Table 2.1 – Recession Duration Data (measured in quarters)

	All Observations	Without Outliers
Mean	7.05	6.58
Median	5	5
Maximum	52	21
Minimum	2	2
Standard Deviation	5.95	4.25
N	148	146

Censored observations are those that have either an unknown beginning time prior to the observation period or an unknown ending time after the observation period. Figure 2.1, based on a figure from Kiefer (1988), illustrates this point. In the sample dataset for this paper, recession A has a starting time of 1979:Q1 because that is the earliest time in the observation period. However, there is no easy way of knowing whether the recession did actually begin in 1979:Q1 or if it began in some quarter prior to that. Recessions B and C are not censored, but recession D is censored since it ends at some unknown point beyond the observation period. Censoring is a problem that is usually unavoidable in duration analysis. The estimation accounts for the fact that these observations are at least the observed length *t* but not equal to it.

Figure 2.1 – Duration Data



The econometric methodology utilized in this chapter is based on the duration analysis research of Kiefer (1988) and Greene (1993, p. 715). In general, duration analysis is very useful for studying the lengths of certain events or the probability that an event will occur given that it has not already occurred. The literature in economics has drawn from that of other fields where duration data have been used to study such things as the useful lives of electronic components, the survival times of organ transplant recipients, and the probability of natural disasters occurring. In economics, duration analysis has most commonly been applied to research on the lengths of unemployment spells.

The key concept in duration analysis is that this statistical method does *not* involve the unconditional probability of an event taking place (e.g., the probability of an unemployment spell ending at exactly six weeks independent of all other time periods) but rather the conditional probability of an event happening (e.g., the probability of an unemployment spell ending at six weeks *given* that it did not end at five weeks). Unconditional probabilities are the emphasis when specification is in terms of probability distributions, but the "hazard function" specification explained below emphasizes the conditional probabilities (Kiefer 1988). As Kiefer points out, individuals tend to reason in terms of conditional probabilities anyway, so duration analysis better links theory to econometric estimation.

Duration analysis is also more useful to policymakers. A common classical criticism of the use of policy is that its results are too slow to take effect and often take effect after the business cycle has already turned. To avoid this, it would be more helpful

if policymakers knew the conditional probability of a recession ending after a certain number of quarters given that it did not end after the previous quarter. In other words, the probability of a recession ending varies over the length of the recession and may increase or decrease depending on policy changes in previous quarters. For example, if a particular state is in the fourth quarter of a recession, it would be helpful to know whether the recession is more likely or less likely to end next quarter. If policymakers know the recession is more likely to end, they can take very small policy measures or none at all. In contrast, if they know the recession is less likely to end, they can enact stronger policy measures in an attempt to reverse the business cycle. Such conditional probabilities are found through duration analysis.

A. Nonparametric Analysis

For the purposes of this paper, duration analysis basically involves estimating the conditional probability that a recession will end in period t given that it has not yet ended in period t-1. The conditional probability function, or the hazard function, is defined as

$$\lambda(t) = f(t) / [1-F(t)] \qquad (1)$$

where $F(t) = \Pr(T \leq t)$ is the cumulative distribution function and $f(t) = dF(t)/dt$ is the corresponding density function (Kiefer 1988). T is a random variable denoting duration and t can be viewed as the realization of that variable. Thus $F(t)$ is the probability of a duration T ending at time t. Then $f(t)$ is the density function which is everywhere non-negative and can be thought of in terms of a histogram. In other words, it depicts the frequencies of empirically measured values of T. Basically, the hazard function is the

probability of a recession ending at time t, denoted by $f(t)$ in the equation, over the probability of the recession lasting until time t, denoted by $[1-F(t)]$. This hazard function is useful for addressing the key question policymakers would like answered, that is, given that a recession has lasted until time t, what is the probability that it will end in the next interval of time, in this case the next quarter?

Nonparametric graphical analysis of the hazard function is not always clearly interpreted, but the general slope of the function can sometimes be seen. If $d\lambda(t)/dt > 0$, then the function is upward sloping and indicates positive duration dependence, meaning the probability of a recession ending increases as the length of the recession increases. As shown later, this is true for the state recession data used in this paper. A downward sloping hazard function would have indicated negative duration dependence, meaning the probability of a state recession ending decreases as the recession lengthens.

In addition to the hazard function, it is also useful to define and examine the survivor function. The survivor function is the probability of a recession lasting or "surviving" until a certain quarter *given* that it did not end in the previous quarter. In other words, the survivor function indicates the probability that the length of a recession, T, will equal or exceed the value t. It is defined as:

$$S(t) = 1 - F(t) \quad \text{or} \quad S(t) = \Pr(T \geq t) \tag{2}$$

Notice that since the survivor function equals 1-F(t), the hazard function from equation (1) above can be written as a function of the density function and the survivor function:

$$\lambda(t) = f(t) / S(t) \tag{3}$$

From this form, the conceptual definition is more easily seen. The hazard function is simply the probability of a spell ending over the probability of a spell "surviving." Basically, it is the rate at which spells end after duration t, given that they last at least until t.

The graph of the survivor function is interpreted in the opposite way of the hazard function. A downward sloping survivor function indicates positive duration dependence while an upward sloping survivor function indicates negative duration dependence. As shown later, the survivor function for the state recession data is downward sloping, meaning a recession is less likely to "survive" the longer it lasts.

In order to plot the hazard and survivor functions on a graph, a sample estimator, $\hat{\lambda}(t)$, is needed. It is constructed in the following way. First, the recessions (of sample size n) are ordered from shortest duration to longest duration, $t_1 < t_2 < t_3 \ldots < t_k$. The number of completed durations k is usually smaller than n due to some observations having the same duration length. For example, there are 17 recessions in this dataset that lasted for three quarters. For this particular dataset, n equals 148 and k equals 23.

Then let h_j be the number of recessions that ended before duration t_j, for $j = 1,\ldots,k$. Let n_j be the number of recessions that did *not* end before duration t_j:

$$n_j = \sum_{i \geq j}^{k} h_i \tag{4}$$

Thus, a convenient sample estimator for $\hat{\lambda}(t)$ is:

$$\hat{\lambda}(t) = h_j / n_j \tag{5}$$

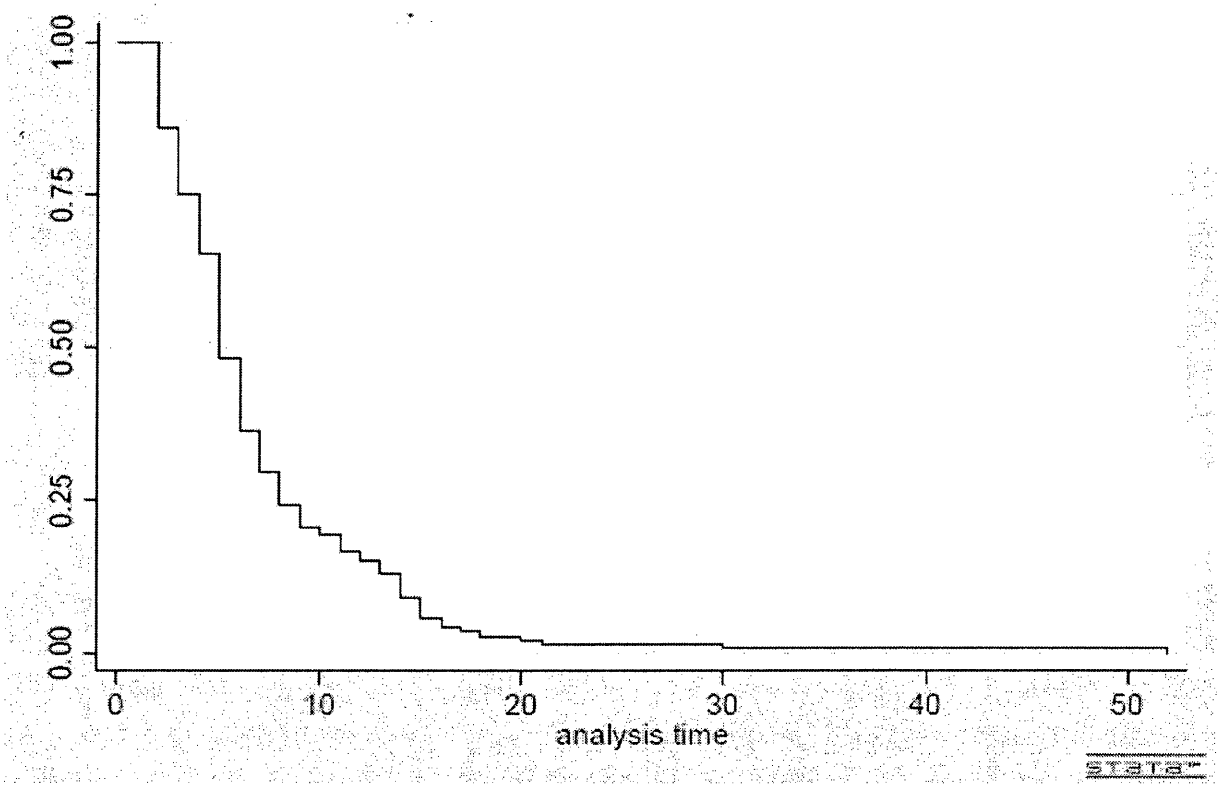
This shows that the estimated sample hazard is simply the number of recession completions at t_j over the number of recession "survivors" at t_j. The corresponding survivor function is then:

$$\hat{S}(t_j) = \prod_{i=1}^{j} (1 - \hat{\lambda}_i) \tag{6}$$

Following (Kiefer 1988), this survivor estimator is obtained by setting the estimated conditional probability of a recession ending at t_j equal to the observed relative frequency of recessions ending at t_j.

The hazard and survivor estimates calculated are presented in Table A2 of the Appendix. The graph of the survivor function is shown in Figure 2.2 on the next page. In this figure, the units on the horizontal axis are quarters since the unit of time for this dataset is quarters, and the units on the vertical axis are the survivor estimates or the probabilities of a recession lasting until time t_j. As explained earlier, the survivor function for this dataset exhibits positive duration dependence since it is downward sloping. This means a recession is less likely to "survive" the longer it lasts.

Figure 2.2 – Survivor Function



Finally, the integrated hazard function is useful for specification checks. It does not have a convenient interpretation since it is not a probability, but it is defined as

$$\Lambda(t) = \int_0^t \lambda(t)\, dt$$

and its relation to the survivor function is

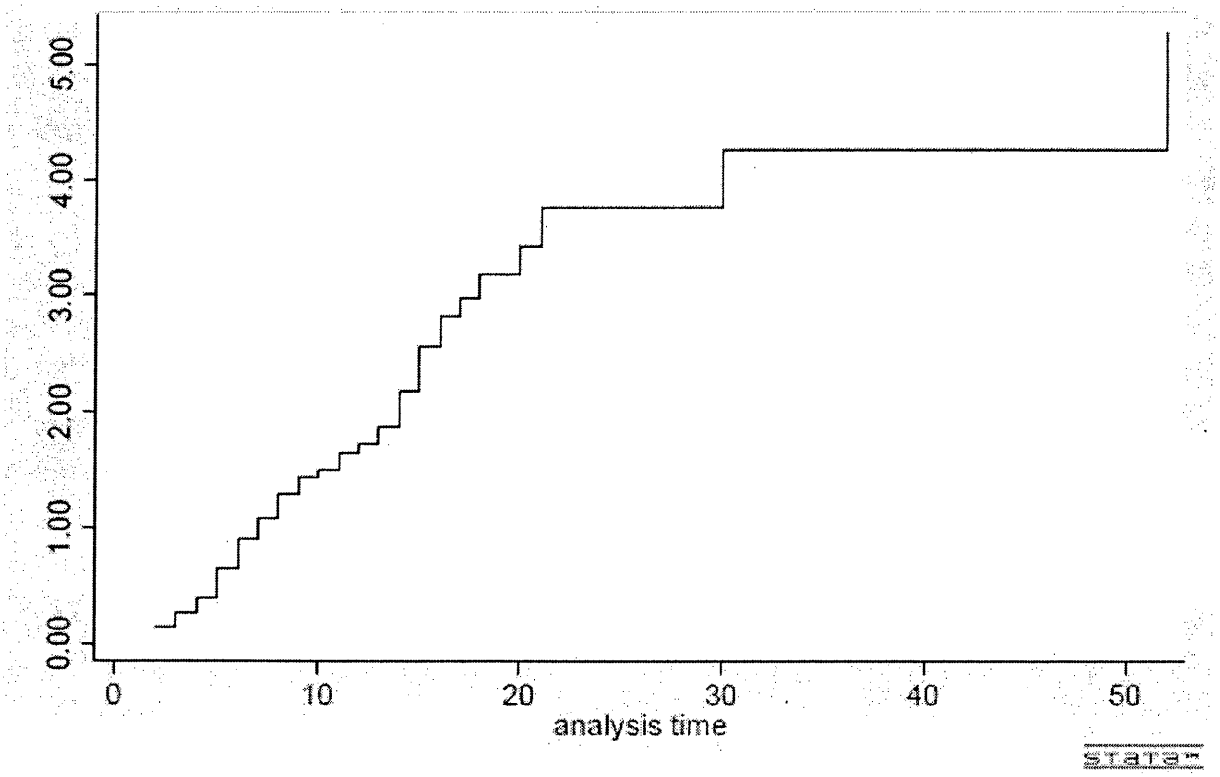
$$\Lambda(t) = -\ln S(t).$$

The integrated hazard estimates for this dataset are also listed in Table A2 of the Appendix, and the graph of the integrated hazard function is shown in Figure 2.3 on the following page. This function appears to be increasing at an *increasing* rate over most of

the range, so it exhibits positive duration dependence as the hazard and survivor functions both indicated previously. Again, this means that the longer a recession lasts, the more likely it is to end soon.

However, near the end of the integrated hazard function, the slope is still positive but appears to be increasing at a *decreasing* rate. This means that after a certain point, the longer a recession lasts, the more likely it is to continue rather than end. Note that the two very long recessions at the right end of the graph are the two censored observations for Hawaii and Alaska that were pointed out earlier.

Figure 2.3 – Integrated Hazard Function



As mentioned earlier, some valuable insight can be gained from simply examining the graphical representations of these functions. For example, the message to be taken by state policymakers from the integrated hazard function shown here is that most recessions are more likely to end as they get longer. However, if a recession continues past a certain point, approximately 20 quarters, drastic measures may be needed in order to reverse the business cycle, otherwise the recession will be more likely to persist rather than end.

While this information is certainly beneficial, it would be much more complete if characteristics of the state economies could be included as parameters in the model. This would give a better understanding of what factors, other than duration, influence the conditional probability that a recession will end. Therefore, policymakers can focus their efforts on changing those factors that most significantly contribute to ending an economic recession.

B. Full Parametric Analysis

Any number of factors could affect the duration of a recession in a particular state. The factor of particular interest in this research effort is industrial diversity. As mentioned earlier, traditional Ricardian trade theory suggests that a state should specialize in one or more industries in which it has a comparative advantage. However, this leaves a state susceptible to developments in those industries that occur outside the state and therefore outside its control. For example, the price of energy and other inputs may change, environmental policies may be tightened, or new innovations may reduce

the competitiveness of a particular industry. Therefore, greater industrial diversity, rather than specialization, would theoretically help insulate a state from the economic influences outside of its control by giving it a wider array of industries to rely upon for stabilizing demand and employment.

The following general model includes industrial diversity along with other factors that are expected to influence the duration of recessions.

$$\text{DURATION} = f\,(\text{DIV, DIVCH, U, RPICH, NWT, TEEN, OVER65, POP, POPCH, DENS})$$

As mentioned previously, the selection of these independent variables is guided by the model presented in Izraeli and Murphy (2003).

DURATION is the length of time over which a state recession lasts, measured as the number of quarters. DIV is the level of industrial diversity in a state in the year prior to the start of the recession. Industrial diversity is measured by the Herfindahl index, which is defined as $\Sigma(E_{is}/E_s)^2$, where E_{is}/E_s is the employment share of industry i in state s. These indexes are calculated from the employment shares for each industry by state over time and these data are provided by County Business Patterns, the Census Bureau's annual report on business activity. Table A1 in the appendix lists the full Web site addresses of all sources. A higher Herfindahl index indicates a less industrially diverse state while a lower index indicates greater industrial diversity. A priori, it is expected that states with greater industrially diversity can better weather an economic downturn and therefore experience shorter recessions.

DIVCH is the change in industrial diversity from the year prior to the start of the recession to the year in which the recession occurred.[5] The change in diversity in addition to the level of diversity is included because the two variables capture different effects and both could have important influences on recession duration. It is hypothesized that increases in industrial diversity should have a negative effect on the duration of recessions. Since the Herfindahl index decreases when diversity increases, a state that experiences a decrease in its Herfindahl index will experience shorter recessions. Thus, a positive sign on DIVCH is expected.

U is the state unemployment rate in the year prior to the start of the recession and comes from the Bureau of Labor Statistics. The one-year lagged value for unemployment is necessary because recession duration, the dependent variable, is constructed using the current unemployment rate. A higher state unemployment rate is expected to be positively associated with longer recessions. Due to the immediate costs of migrating in the short run, unemployed workers are more likely to stay where they are rather than to move to find employment. This keeps the unemployment rate higher and contributes to longer recessions.

RPICH is the change in real per capita income at the state level from the year prior to the start of the recession to the year in which the recession occurred. The data for this variable come from the Regional Economic Information System (REIS), a subdivision of the Bureau of Economic Analysis. The change in state income is expected

[5] For recessions that spanned more than one year, the values for those variables are averages of all years during which the recession lasted.

to be negatively related to recession duration as increasing incomes would be more likely to lead to shorter recessions while decreasing incomes would likely lengthen recessions. The change in income rather than the income level is used because it is more consistent with how recessions are defined. At any given point in time, the income level varies greatly among states, but this does not necessarily determine which states are experiencing recessions versus expansions. The measure that matters more is whether income is increasing or decreasing.

Following Izraeli and Murphy (2003), three demographic variables are included in the model to account for different population characteristics among states. All three are calculated from U.S. Census data. NWT is the percentage of the working age population that is non-white. More specifically, this is calculated as the percentage of 15- to 64-year-olds who are of any race other than white. This variable is expected to be positively associated with recession duration since this group tends to experience higher unemployment rates (Izraeli and Murphy 2003).

The expected effect of TEEN on recession duration is unclear. TEEN is calculated as the percentage of 15- to 64-year-olds who are 15 to 19 years of age. Like non-white workers, teenage workers tend to experience higher unemployment rates than the general population, so TEEN could be positively related to duration. On the other hand, a larger proportion of teenage workers means a smaller labor force, which should lead to a lower unemployment rate and thus a shorter recession.

The third demographic variable, OVER65, is included to represent the proportion of the total population that is 65 years of age or older. The a priori expectation for its

effect on duration is also unclear. The per capita income and consumption level of this population group is relatively low, so this could contribute to longer recessions given the weaker demand. On the contrary, a higher proportion of retired individuals means a smaller labor force, which would likely cause lower unemployment and shorter recessions.

Finally, measures of state population, population change, and population density are added to the model. POP is the log of total population in each state in the year prior to the start of the recession. The data for total population are reported by REIS, and this variable could affect the length of recessions in different ways. A higher population could be an indicator of economies of scale, making a state's businesses more competitive and thus its recessions shorter in duration. In contrast, population could be associated with a higher cost of living and/or more generous welfare payments, which would contribute to longer recessions.

POPCH is simply calculated from population as the log difference in a state's total population from the year prior to the start of the recession to the year in which the recession occurred. It is expected to be negatively associated with recession duration because the in-migration of people tends to increase the proportion of a state's labor force that is likely to be employed.

DENS represents population density or the average number of persons per square mile in a state. It indirectly reflects production costs because a higher population density can lower transportation, communication, and labor costs, for example. If production costs are lower, that state's industries can be more competitive and contribute to shorter

recessions. Conversely, areas with sparse populations are more likely to have higher production costs as well as fewer businesses to absorb laid-off workers. So a priori, the sign on DENS is expected to be negative, meaning a higher population density is associated with shorter recessions.

Table 2.2 – Descriptive Statistics

Variable	Mean	Std. Deviation	Minimum	Maximum
DURATION	7.08	6.00	2.00	52.00
DIV	318.00	47.14	236.64	470.77
DIVCH	0.03	0.05	-0.08	0.36
U	5.96	1.83	2.22	12.43
RPICH	-0.01	0.03	-0.17	0.04
NWT	12.74	10.51	0.74	63.95
TEEN	12.72	1.70	9.13	16.10
OVER65	11.54	2.10	2.83	18.27
POP	4,701,909	4,886,951	402,191	29,218,165
POPCH	0.02	0.02	-0.01	0.11
DENS	170.41	239.95	0.71	1050.85
N	*145*			

More detailed descriptions of the data sources for each variable are listed in Table A1 of the Appendix. Descriptive statistics and correlations for these variables are

provided in Tables 2.2 and 2.3, respectively. In Table 2.2, POP is reported in terms of total population levels. In the analysis, the natural log form of total population is used.

The sample size equals 145 recessions instead of the original 148 because when examining a scatterplot of duration against diversity, three observations appear to be skewing the results. Nevada's industry is very highly concentrated and thus has a much higher Herfindahl index relative to the other 49 states. While Nevada's Herfindahl index values are very high, their recession durations are relatively short and it is clear that these three observations are outliers. Therefore, Nevada's three recessions are dropped from the dataset for the full parametric analysis, leaving 145 state-level recessions. Note that these 145 observations represent all state-level recessions that occurred between 1979:Q1 and 1997:Q4 regardless of the state. In other words, there is no set number of recessions per state. One state may have had only one recession during this time period while another state may have had five, for example.

Table 2.3 - Correlations

	DURATION	DIV	DIVCH	U	RPICH	NWT	TEEN	OVER65	POP	POPCH	DENS
DURATION	1.000										
DIV	0.095	1.000									
DIVCH	0.179	-0.185	1.000								
U	0.009	-0.057	-0.161	1.000							
RPICH	-0.084	-0.147	0.188	0.085	1.000						
NWT	0.217	-0.104	0.003	0.117	0.047	1.000					
TEEN	-0.158	-0.147	-0.309	0.234	-0.196	-0.102	1.000				
OVER65	-0.194	0.118	0.123	-0.279	0.150	-0.231	-0.226	1.000			
POP	-0.194	-0.481	0.001	0.047	0.127	0.248	-0.090	0.175	1.000		
POPCH	0.539	-0.015	0.071	-0.131	-0.076	0.189	-0.187	-0.391	-0.064	1.000	
DENS	-0.023	-0.186	0.115	-0.080	0.178	0.080	-0.157	0.243	0.246	-0.224	1.000

In order to utilize full parametric maximum-likelihood estimation and interpret the estimated coefficients, a probability distribution must be chosen. The hazard function still has an equivalent specification in terms of a probability distribution, but since the normal and lognormal distributions do not allow for a constant hazard, the exponential, Weibull, and log-logistic distributions are more commonly used in duration data analysis (Kiefer 1988). The corresponding hazard function for each of these three common distributions is listed in Table 2.4.

Table 2.4 – Principal Hazard Distributions

Distribution	Hazard Function, $\lambda(t)$
Exponential	λ
Weibull	$\lambda p(\lambda t)^{p-1}$
Log-logistic	$\lambda(t) = \lambda p(\lambda t)^{p-1} / [1+(\lambda t)^{p}]$

Source: Greene (1993, p. 718)

The hazard function of the exponential distribution is constant over time and therefore reflects no duration dependence. It depends on only one parameter, λ, so it is simple to work with and interpret. However, because the exponential distribution depends on only one parameter, the mean and variance cannot be adjusted separately and therefore this distribution is not likely to be adequate if the dataset has a lot of variation in the duration lengths (Kiefer 1988).

When the hazard probability varies over time, the Weibull distribution is more appropriate. The hazard function of the Weibull distribution is monotonically increasing or decreasing and depends on two parameters, λ reflecting explanatory variables and p reflecting duration dependence. Duration dependence does not depend on the value of λ but does depend on the value of p. If $p > 1$, the hazard increases. If $p < 1$, it decreases. Note that if $p = 1$, the Weibull and exponential distributions are the same.

Finally, the hazard function for a log-logistic distribution is nonmonotonic with parameters $\lambda > 0$ and $p > 0$. If $p > 1$, the hazard first increases with duration, then decreases. If $0 < p \leq 1$, the hazard decreases with duration.

For each distribution, the explanatory variables enter the model through λ. Models are estimated for each of the three distributions and presented in the following section. When analyzing the results, more attention is given to the Weibull distribution because it is the most commonly used distribution for duration analysis and because graphically, it most closely matches the integrated hazard function in Figure 2.3.

4. Estimation Results

The STATA 8.0 statistical software program is used to perform the analyses. STATA has the capability to appropriately account for the censoring issue explained earlier. The regressions have been corrected for heteroskedasticity as well using White's correction. The results are presented in Table 2.5. The signs and significance are nearly identical between the results for the exponential and Weibull distributions, but they are

somewhat different for the log-logistic distribution. The table lists each coefficient with its t-statistic in parentheses.[6]

The joint hypothesis that all coefficients are jointly equal to zero can be rejected because the Wald χ^2 statistic, indicating goodness of fit, is statistically significant. In other words, a significant Wald χ^2 statistic shows that the coefficients on the independent variables are *not* all jointly equal to zero. This discussion of results focuses on the Weibull distribution for reasons explained in the prior section and because the Wald χ^2 is the highest for the Weibull distribution. In addition, the p value of 1.86 is statistically significantly different from one, indicating that the Weibull distribution is more appropriate than the exponential.

[6] For comparison, ordinary least squares regression analysis was performed on these data also and the results regarding diversity were the same. That is, DIVCH was positive and significant while DIV was not statistically significant.

Table 2.5 – Full Parametric Maximum-Likelihood Estimation Results

Parameter	Exponential	Weibull	Log-logistic
DIV	0.00	0.01	0.00
	(1.01)	(0.79)	(0.81)
DIVCH	3.57**	3.60**	2.78*
	(2.13)	(2.17)	(1.65)
U	0.06*	0.05	0.06*
	(1.66)	(1.48)	(1.90)
RPICH	-3.86*	-4.34**	-2.68
	(-1.74)	(-2.15)	(-1.13)
NWT	0.01*	0.01**	0.01
	(1.86)	(2.26)	(0.95)
TEEN	0.01	0.01	0.00
	(0.20)	(0.37)	(0.03)
OVER65	0.05	0.05	0.03
	(1.52)	(1.53)	(0.78)
POP	-0.14**	-0.15***	-0.10
	(-2.15)	(-2.53)	(-1.26)
POPCH	15.71***	13.32***	16.43***
	(4.83)	(4.81)	(5.75)
DENS	0.00	0.00	0.00
	(1.22)	(1.22)	(1.50)
N = 142			
Wald χ^2	45.84***	64.45***	51.74***
shape parameter		$p = 1.86$	$\gamma = 0.34$

*** *indicates significance at the 1% level*
** *indicates significance at the 5% level*
* *indicates significance at the 10% level*

As expected, the change in diversity (DIVCH) is positively associated with recession length. Recall that the Herfindahl index is higher when industrial concentration is higher. Therefore these results indicate that when industrial concentration is increasing over the year prior to the start of a recession, the duration of the recession is longer. Conversely, when industrial diversity is increasing, recession durations are shorter. Interestingly, the level of diversity (DIV) was not a significant factor in determining recession length. Even if the change in diversity variable is removed from the model, the level of diversity does not turn out to be significant. So according to these results, it is not the initial level of industrial diversity that matters but whether industrial diversity is increasing or decreasing.

At first glance, this result for the level of diversity may appear to be out of line with previous research. For example, the Izraeli and Murphy (2003) study, which was the inspiration for this research effort, reports that the level of diversity reduces unemployment. With unemployment being a key ingredient of a recession, it seems logical that the level of diversity should thus reduce the duration of recessions, but it is an insignificant influence according to the model in this chapter.

However, there is an important difference between this study and previous research, including Izraeli and Murphy. Most previous research examines one continuous time period of business cycles while this chapter isolates the recessionary periods only. There is evidence that economic agents behave differently in recessions versus expansions. For instance, unemployment during expansions is dominated by new entrants and reentrants into the labor market while unemployment during recessions is

dominated by lay-offs (for this and other examples, see Chapters 5 and 6 of Davis et al., 1996). In addition, the Owyang et al. (2004) study reviewed in Section 2 finds that industrial mix affects growth rates differently depending on whether they examine recessions or expansions. Thus, the finding that diversity level is an insignificant influence may stem from which time periods are included in the analysis. The result may be different for expansionary time periods.

Other variables were also found to be significant determinants of recession duration. The change in real per capita income (RPICH) is statistically significant and is negatively associated with duration. In other words, when real income is increasing, recessions are shorter in length. This result is expected since the change in income is an indicator commonly used to define and date changes in the business cycle. If incomes are increasing, then an economy has most likely reached its turning point in a recession.

Another significant determinant of recession length is the percentage of the working-age population that is non-white (NWT). As expected, it is positively associated with longer recessions. The non-white population tends to experience higher levels of unemployment, as Izraeli and Murphy (2003) find in their study also, and this contributes to longer recessions.

Finally, both population (POP) and population change (POPCH) are statistically significant factors, but their coefficients have opposite signs. The opposite direction of their signs is an interesting result. A large total population in the year prior to the start of a recession is associated with shorter recessions, but when the state's population is increasing, recessions are longer. The sign on total population was expected since a

larger population may be an indication of economies of scale in a state, making its businesses more competitive. The sign on population change was not expected since a growing population is sometimes seen as a sign of a thriving economy. It is possible though that a state might not have enough jobs to accommodate in-migration and thus increases in population contribute to longer recessions. The signs on total population and population change are opposite of what Izraeli and Murphy (2004) report, but again, this can likely be attributed to the fact that recessions are isolated here where all recessions and expansions were used in their study. Furthermore, when one continuous time period is examined, expansions comprise a larger proportion of time than do recessions.

The one-year lagged unemployment (U) is the only variable that turns out to be significant in regressions based on the exponential and log-logistic distributions but insignificant in the regression based on the Weibull distribution. The positive sign on unemployment was in line with a priori expectations. A higher unemployment rate is associated with a longer recession. Unemployment, like real per capita income, is another indicator that is commonly used to define and date changes in the business cycle. If unemployment is low, an economy is most likely near the end of its recession.

In general, these results follow the a priori expectations. The sign on population change is one exception. The results also generally support the idea that diversification is a reasonable goal because increased diversification appears to be associated with shorter recessions.

5. Conclusions and Policy Implications

The goals of this paper were (1) to test whether state-level recessions are more likely or less likely to persist as they increase in duration, and (2) to discover what factors and characteristics of state economies influence the duration of recessions, testing the influence of industrial diversity in particular. Regarding the first goal, the main finding is that most recessions are *more* likely to end as their duration increases given that they have not ended already. Research for the second goal revealed that numerous factors have statistically significant influences on the duration of recessions. These factors include the change in industrial diversity, unemployment, the change in real income per capita, the proportion of non-white workers, total population, and the change in total population.

The most interesting and noteworthy finding of this study is that the initial level of industrial diversity in a state is not what significantly influences recession duration. Rather it is whether or not industrial diversity is increasing. In states where diversity is increasing, recessions are shorter in duration. In states where diversity is decreasing, recessions are longer in duration. This research finding is a hopeful message for states that are relatively more concentrated since it seems to indicate that their initial level of diversity is not as important as their change in diversity. Therefore, this could mean that as long as a state can accelerate its industrial diversification, they may be able to reduce the duration of their recessions.

Precisely how to increase diversity, though, is a challenging policy dilemma. Aiming to decrease a state's most prominent industry in order to even out industry

proportions is of course, not very logical. Nor is it logical to attempt to foster industries for which a state does not already have some related potential and capabilities. As a simplistic example, a state should not try to increase its mining industry if it does not already have minerals abundantly available. There may be a good reason why some industries are very small in proportion. Each state should assess its strengths and capabilities when deciding which industries to target for expansion. Similarly, each state should identify which industries are most likely to contract, assuming that not all of the new employment in the expanding industries will come from in-migration. Then, in order to ease the reallocation of labor, policy initiatives such as job training could be pursued.

Before promoting active policy intervention however, it is important to recall the results of the non-parametric analysis in context with the classical argument against the use of policy. The non-parametric analysis suggests that a recession is more likely to end the longer it persists. Therefore, a state should evaluate how long their recession has already lasted and consider whether policy intervention is necessary.

Targeting diversification though, does not appear to have a downside in terms of real per capita income tradeoffs, according to Izraeli and Murphy (2003). This finding, in addition to the general view that diversification reduces instability (Wundt 1992), suggests that a policy aimed at diversifying a state's industrial base is not likely to have negative repercussions even if the recession should happen to end before the policy takes effect. This cannot be said with certainty however, unless expansionary periods are researched as well.

Thus an important direction for future research, as alluded to in Section 4, would be to perform a similar duration analysis on expansionary time periods rather than recessionary periods. If job creation and destruction vary considerably over the business cycle as Davis et al. (1996, p. 83) suggests, then industrial diversity may have different influences over the business cycle as well. Note that an entirely different model may be necessary as determinants of expansion duration may differ from determinants of recession duration. Even so, it would still be interesting to see a comparison between recessions and expansions in terms of how diversity influences the duration of the downturns and upswings in the economy.[7]

To continue the theme of industrial diversity and its influence on regional economic performance, the next chapter addresses small firm formation at the state level. Industrial diversity along with several other factors are included as potential determinants of new firm formation rates. Just as a better understanding of recession duration is useful to policymakers, a fuller understanding of the motivating forces behind small firm formation can be valuable as well.

[7] Another important consideration for future research would be to control for regional spillovers among states. It is possible that recession duration is influenced by the business cycles in neighboring states.

III. Determinants of Small Firm Formation

1. Background

Two general shifts in public policy regarding business have been occurring over the last decade or so. From a federal perspective, government has become more concerned with promoting the startup of small firms and less preoccupied with constraining the large existing corporations (Gilbert et al. 2004). From the state perspective, policies are now more commonly aimed at fostering new entrepreneurs as a source of job growth rather than attracting new branch plants from established businesses (Henderson 2002). Regardless of the level of government, the use of entrepreneurship policy is growing and we need to gain a better understanding of exactly what factors most influence entrepreneurship. In a 1999 study, Georgellis and Wall remark, "Given the extent of the regional variation in entrepreneurship, it is surprising that so little attention has been paid to its determinants" (p. 3). The purpose of this chapter is to examine which regional economic and socioeconomic factors motivate a person to become an entrepreneur by starting a new business. As explained on the next page, being an entrepreneur and starting a new business are not exactly the same, but new firm formation can serve as a readily available proxy for entrepreneurship.

This topic has been approached from other angles as well and a wide body of literature from various disciplines exists providing valuable insight into the determinants of success for an individual entrepreneur, such as personality characteristics, organizational structures, and management practices. However, a gap remains in our

understanding of the larger context that fosters entrepreneurs. Why do some states and regions have very large, thriving entrepreneurial bases while others do not? What characteristics of regional environments most influence entrepreneurs? In this regard, economics and economic modeling may fill the void that other disciplines have not yet completely filled.

A clear definition of entrepreneurship would make the study of it much more straightforward. However, a consensus on its appropriate definition is very difficult to reach. Malecki (1994) explains that entrepreneurship occurs on three general levels. At the most basic level, entrepreneurship refers to any informal economic activities, including black market and underground, that occur outside the recognized and fully legal economic activities. Entrepreneurship at the next higher level refers to new small firm formation or any new enterprises added to the economy. At the highest level, entrepreneurship takes on the Schumpeterian view that innovation is the key characteristic of an entrepreneur. This definition means that entrepreneurship does not necessarily require the formation of a new business, rather it entails the creation of new products, processes, or markets that then foster new firm formation.

For the purposes of this chapter, it is assumed that entrepreneurship is nearly synonymous with small firm formation, following Malecki's middle level definition. This appears to be the most common interpretation of the term when used in policy discussions and the media. Further, data on small firm formation are readily available while data measuring Malecki's other two forms of entrepreneurship are much more difficult to obtain.

Studying small business creation is important for various reasons. Numerous studies and statistics show that a majority of new jobs are created by new start-ups (Armstrong and Taylor 2000, p. 264). New firms add to both employment and output to boost a state's economic performance. They also contribute to a more flexible and diversified labor market. Moreover, evidence shows that new firms stimulate competition with existing businesses, and even more importantly, they stimulate innovation. While innovation-based new firms may represent a small proportion of start-ups, they often generate entirely new industries based on their innovations in products or processes, further challenging established businesses to grow and improve.

New firm formation is not the cure-all for economic woes, of course, as employment by start-ups has its disadvantages. While some would argue that small firms offer superior working environments, others often criticize small firms for offering low-skilled and part-time positions that pay lower wages with less training and fewer benefits. In general, employment by start-ups is less stable due to the high volatility of small firms. For example, only 55 percent of newly created firms survive three years (Storey 1994). Nonetheless, as the trend continues to shift toward promoting small business formation, research is needed to create policies that promote the survival of small businesses as well as their initial formation.

The rest of this chapter is organized as follows. Relevant literature is reviewed in the next section. Section 3 outlines the model and explains the independent variables. The results are presented and analyzed in Section 4. The final section elaborates on the conclusions from the model and discusses some policy implications.

2. Literature Review

Over the past several years, researchers have been increasingly interested in the regional determinants of new firm growth, yet there seems to be little consensus on what factors are most important. This makes it confusing for policymakers looking for guidance in drafting regional policy. The lack of consistent results likely stems from the fact that entrepreneurship is difficult to measure and therefore different units of analysis are used. Further, there are numerous possible independent variables that could be used, adding to the difficulty of drawing comparisons between results.

There is one strand of literature, particularly studies on British data, which focuses on self-employment as the measure of entrepreneurship. They argue that self-employment is a suitable proxy for those who have started their own business, especially when more detailed data on new firm formation are unavailable. Georgellis and Wall's 2000 study is an example. They examine self-employment rates across regions in Britain from 1983 to 1993, and they use numerous explanatory variables to capture four main influences: labor market conditions, labor force characteristics, industry composition, and region-specific effects. They conclude that all four have a significant impact on self-employment rates, especially the labor force characteristics like age, gender, and education level.

Another larger thread of literature relies on actual firm birth rates, sometimes weighted by the population or labor force, to measure entrepreneurship rather than self-employment rates or growth rates of small firms. Armington and Acs (2002) and Lee et al. (2004) use actual firm births per 1,000 population for 394 Labor Market Areas in the

U.S. between 1994 and 1996. Both studies find population growth, income growth, industry density, and human capital to be positively associated with new firm formation, while mean establishment size is negative. Sutaria and Hicks (2004) focus on Texas metropolitan areas only and some of their results contradict those of the previous two papers mentioned. They find that population growth and income growth have no significant effect, while mean existing establishment size is positively associated with new firm formation. Sutaria and Hicks also find that greater availability of financial capital is positively associated with new firm formation.

Reynolds, et al. (1995) use actual firm births as the measure of entrepreneurship also, but they expand the scope of their study to include cross-national comparisons. They examine new firm formation in six countries during the 1980s and conclude that regional variations within countries are roughly similar. For the United States in particular, they examine 382 regions and find that the statistically significant positive determinants of firm births include population growth, GDP growth, percentage of managers in the work force, the unemployment level, dwelling prices, and industry specialization. Local government expenditures and the percentage of workers with higher education were negatively associated with new firm formation.

Using firm births as the dependent variable has at least one major drawback in that it does not account for the size of the start-up. Many new firms either do not survive over the long run or do not increase their output or employment significantly enough to impact a state's economic performance. Isolating high-growth entrepreneurs would offer more insight, but this is difficult to define. One suggestion is to use rates of incorporation

instead of simply rates of new firm formation (Henderson 2002).[8] Evidence shows that many people start out as self-employed then incorporate their businesses after they begin to grow. Incorporated entrepreneurs tend to have higher incomes and larger firms, possibly indicating more growth potential than unincorporated entrepreneurs.

Wall (2004) is one example of a study using incorporations as the dependent variable. Wall is specifically interested in investigating the effect of banking deregulation on entrepreneurship rates, so all of his explanatory variables are related to banking. His general conclusion was that the deregulation of banking had no consistent effect upon entrepreneurship rates.

A unique approach in the literature on entrepreneurship is to isolate the most successful new small firms based on growth rates and examine the determinants of their prevalence in certain regions. Friedman (1995) adopts this approach, focusing on newly established small firms that were identified as "top" firms by the popular business publications of *INC.*, *Business Week*, and *Forbes*. She examines 208 US urban areas in the 1980s and concludes that the major determinants of the distribution of high-growth entrepreneurs are total population, percentage change in employment, percent college graduates, amenities, industrial diversity, and the number of venture capital firms. All of these are positively associated with the prevalence of "top" new firms. While these results are quite interesting, relying on magazines to identify high-growth start-ups is not necessarily a very comprehensive measure for all US regions.

[8] The initial intent of this analysis was to use incorporations per state but both the Small Business Administration and Dun & Bradstreet stopped reporting this data as of 1998.

Another way to focus on high-growth entrepreneurs is to disaggregate the firm formation data based on size. This of course has drawbacks as well. For example, starting out with 100 or more employees does not necessarily guarantee that a new firm will succeed and experience high growth. Similarly, a new firm that has fewer than twenty employees in its initial year may expand output and employment rapidly in its subsequent years. While acknowledging that initial firm size is not perfectly related to growth potential, this research separates the data by initial start-up size in an effort to capture the size aspect. A simple dummy variable for size could have been added to determine whether size matters, but splitting the sample has the advantage of showing how each independent variable affects small firms in particular.

The model in this thesis is an attempt to synthesize the strong points of the models reviewed above as well as to build on them. This research effort is different from previous studies in three main ways. First, it includes some notable variables of interest, such as environmental hazardous waste sites and Internet access, that have not previously been included in the literature on small business start-ups but would logically be expected to affect business start-up decisions.

Second, this model includes more comprehensive measures of certain independent variables than previous studies. For example, the state industrial Herfindahl index is included as a possible determinant of small business creation in line with the theme of this thesis on the influences of industrial diversity. Studying industrial diversity is interesting and important because existing theories are conflicting. Traditional Ricardian trade theory suggests that a region should specialize in order to grow while other research

indicates that a region should diversify in order to provide more stability and niche markets for growth. Various measures of industrial structure have been examined as determinants of firm formation, but the Herfindahl index specifically has not commonly, if ever, been utilized.

Nor have many studies utilized a broad measure of the overall business environment for start-ups. Many studies have included some measure of tax rates, but the business climate variable used here and explained later is a ranking of numerous and varied public policy aspects, including taxes, that affect the costs and operation of new firms. The inclusion of these more comprehensive measures paints a better picture of the reality that entrepreneurs are facing when deciding whether to start their own business.

Third and finally, this thesis updates the previous literature since it is based on 2001 data. The popularity of this topic of new firm formation peaked during the 1980s and has only recently begun to attract significant attention again from economists (Armington and Acs 2002). Therefore, the data utilized in most regional models of new firm formation are from the 1980s and early 1990s. Arguably, the rapid economic expansion and technology boom of the late 1990s changed underlying trends in the national economy, warranting a fresh look at business start-up behavior.[9].

From the results of regression analysis, five determinants appear to exert significant influence on the rate of new firm formation. Greater industrial diversity, more readily available financing, and larger mean existing establishment sizes are associated

[9] The selection of 2001 data was not a random choice but was simply determined by data availability. 2001 was the most recent year for which new firm formation data were available, yet 2001 was one of the earliest years for which Internet access data were available. It is acknowledged that this may limit the comparability of this study to other studies.

with higher rates of firm formation, while higher education levels and larger numbers of hazardous waste sites are associated with lower rates of new firm formation. These relationships are very similar regardless of whether the model uses small- or medium-sized start-ups as the dependent variable. The most notable result relative to previous studies is the influence of hazardous waste sites on new firm formation.

3. Model and Data

Models of new firm formation rates in the previous literature have used a wide variety of factors to explain differences in the entrepreneurial bases among regions. Storey (1994) provides a useful outline of eight general factors that influence the start-up of new firms and this model will attempt to capture each of those eight influences with various proxy variables. The proposed model is as follows:

$$NEW = f(POPCH, MES, DIV, RPICH, FIN, EDUC, U, CLIM, WEB, ENV)$$

These variables are defined and described in more detail below (see Table 3.2 also).

A. Dependent Variable

The dependent variable, NEW, is the annual total number of new establishments in each state in 2001. These data are collected and reported by the Statistics of U.S. Business (SUSB), a subdivision of the U.S. Census Bureau. 2001 is the most recent year for which data are available. Models are estimated using two different size categories: small firms (those with fewer than 20 employees in their first year of business) and

medium firms (those with 20 to 99 employees in their first year). The respective variable names are thus NEW_SM and NEW_MD. Numerous empirical studies have found that smaller firms tend to grow faster than larger ones (Parker 2004, p. 215). As mentioned earlier though, these size categories do not necessarily reflect growth potential. Growth rates are not the focus of this research effort, but some insight may be gained by distinguishing among new establishments based on their initial start-up size. Of the total new firm establishments in 2001, 77.3 percent were small firms, according to the categories used for this paper, and an additional 4.3 percent fell into the medium category.

Table 3.1 lists the fifty states in order by total new firms in 2001 (after summing the small and medium totals). For comparison, it also lists two proportions: new firms per 1,000 people employed and new firms per 100 existing establishments. The final column in Table 3.1 shows the percentage increase in small and medium establishments due to new firm births in 2001 (this does not take into account firm deaths).

Table 3.1 – New Firm Formation Rates in 2001

State	Total New Firms (small + medium)	New Firms/ 1,000 Employed	New Firms/ 100 Establishments	% Change in Establishments Due to New Firms
California	77,458	6.01	11.00	13.5
Florida	45,218	7.28	12.10	15.1
New York	42,513	5.78	9.80	11.6
Texas	42,218	5.26	9.98	13.0
Illinois	22,196	4.04	8.07	10.1
Pennsylvania	19,876	3.91	7.40	9.4
New Jersey	19,822	5.59	9.60	11.5
Georgia	18,812	5.40	10.61	13.9
North Carolina	17,547	5.18	9.59	12.3
Ohio	17,374	3.47	7.04	9.1
Michigan	17,307	4.25	8.20	10.3
Washington	15,150	6.68	10.54	12.8
Virginia	14,234	4.90	9.02	11.8
Colorado	14,036	7.34	11.88	14.6
Massachusetts	13,812	4.47	8.78	10.8
Arizona	11,264	5.87	11.20	14.6
Missouri	10,788	4.50	8.24	10.4
Minnesota	10,643	4.44	8.72	10.8
Maryland	10,624	5.16	9.31	11.8
Indiana	10,254	3.87	7.70	10.0
Tennessee	9,910	4.15	8.31	11.1
Wisconsin	9,475	3.92	7.49	9.3
Oregon	9,097	6.71	10.25	12.5
South Carolina	8,148	5.09	9.32	12.0
Louisiana	7,781	4.89	8.46	10.8
Alabama	7,684	4.65	8.43	10.9
Oklahoma	7,118	5.93	9.29	11.6
Kentucky	6,637	4.39	8.10	10.5
Connecticut	6,206	4.01	7.45	9.1
Utah	5,974	6.52	12.52	15.9
Kansas	5,780	5.12	8.57	10.8
Nevada	5,450	6.04	13.13	17.1
Arkansas	5,329	5.38	9.42	11.8
Iowa	5,109	4.04	6.96	8.7
Mississippi	4,730	4.94	8.72	11.2
Nebraska	3,737	4.98	8.40	10.4
New Mexico	3,721	6.78	9.72	12.3
Idaho	3,720	8.25	11.48	13.8
Maine	3,204	6.52	9.36	11.2
Montana	2,903	9.80	10.55	12.3
New Hampshire	2,884	5.28	8.65	10.6
West Virginia	2,721	4.88	7.23	9.2
Hawaii	2,380	5.51	8.77	11.0

(continued on following page)

State	Total New Firms (small + medium)	New Firms/ 1,000 Employed	New Firms/ 100 Establishments	% Change in Establishments Due to New Firms
(continued from previous page)				
Delaware	2,091	5.54	9.94	12.9
Rhode Island	2,050	4.94	8.05	9.6
South Dakota	1,885	6.15	8.92	10.7
Wyoming	1,669	9.57	10.51	12.4
Alaska	1,545	7.55	10.10	12.1
Vermont	1,516	5.98	7.91	9.2
North Dakota	1,319	5.17	7.26	8.8

Source: Statistics of U.S. Business.

Noting the differences between the total new firms and the proportions of new firms, the question may arise here as to why the total was used instead of a proportion. There are several reasons. For one, there does not seem to be a general consensus in the literature on which is more appropriate to use, and the more recent of the studies reviewed above utilized the total rather than a proportion (Sutaria and Hicks, 2004, and Wall, 2004). Secondly, imposing a constant and unitary elasticity between the dependent variable and a scale variable may be too restrictive. The inclusion of mean existing establishment size and hazardous waste sites per person, explained below, already controls for scale indirectly. Finally, using the total new establishments simply met the goal of this paper better, that is to target the behavior of new start-ups. Policymakers most commonly seek to increase the total number of establishments, not a proportion based on the number employed or the number of existing establishments.

B. Independent Variables

The potential determinants of new firm creation are numerous, varied, and often intangible, as Malecki (1994) explains. Storey (1994) also points out that the list of factors can be very long, but he summarizes them into the eight general influences listed in Table 3.2. This list of eight guided the selection of independent variables for this model to ensure that all of the general influences based on previous research were accounted for. At least one variable was chosen to represent each category from Storey's list. Table 3.2 lists the proxy variable(s) used in this analysis next to its corresponding category. The ninth row is added to capture the network effects and knowledge access that have become more important and better understood in recent years. As Armington and Acs (2002) mention, there have been new theoretical developments regarding spatial perspectives, agglomeration, localization, and economic growth that have affected our modeling of new firm formation rates. See Table A3 in the Appendix for a complete definition of each variable and its source.

Table 3.2 – Potential Determinants of Small Firm Formation Rates

	Determinant Group (based on Storey 1994)	**Proxy Variable**	**Expected Sign**
(1)	Population and Its Characteristics	POPCH	positive
(2)	Industrial Structure	MES DIV	indeterminate negative
(3)	Wealth/Income	RPICH	positive
(4)	Owner-Occupied Housing (proxy for finance)	FIN	positive
(5)	Occupational/Educational Characteristics	EDUC	positive
(6)	Unemployment	U	indeterminate
(7,8)	Government and Policy Initiatives	ENV CLIM	negative negative
[9]	Network Effects/Access to Knowledge	WEB	positive

Population growth, represented by the variable POPCH, is calculated as the log difference in state population from 1999 to 2000, and it is based on population data available from the Regional Economic Information System (REIS), a subdivision of the Bureau of Economic Analysis. The a priori expectation is that population growth is positively associated with new firm formation because it is thought that growth stimulates business start-ups (Armstrong and Taylor 2000, p. 278). An increase in population leads to both an increase in demand for goods and services as well as an increase in the pool of labor, both of which should encourage the formation of new firms.

Two variables are included to capture the industrial structure characteristics in each state. MES is the mean establishment size or the average size of the existing establishments in each state in 2000, calculated from SUSB data. Studies have found

mixed results for this factor, as mentioned earlier. It is commonly hypothesized that new firm formation is higher in areas where there already exists many small firms because a lower MES indicates an area that has already restructured away from large manufacturing dominance (Armington and Acs 2002). So a higher MES indicates a greater dominance by large firms and therefore MES may be negatively associated with the dependent variable. Alternatively, large firms may actually play a positive role by purchasing inputs from, as well as outsourcing work to, small neighboring firms, suggesting a positive relationship between MES and new firm formation (Sutaria and Hicks 2004).

The other variable representing industrial structure in the model is the industrial diversity factor that is the common theme of this thesis. As in the earlier model on recession duration, DIV is the Herfindahl index for each state in 2000, defined as $\Sigma(E_{is}/E_s)^2$, where E_{is}/E_s is the employment share of industry i in state s. These indexes are calculated from the employment shares for each industry by state over time and these data are provided by County Business Patterns, the Census Bureau's annual report on business activity. A higher Herfindahl index indicates a less industrially diverse state while a lower index indicates greater industrial diversity. The sign on this variable depends on whether new firms can more easily compete when there is a wider variety of industries amongst which to fill niche markets or whether new firms can more easily compete when they are supporting and benefiting from a few large growth industries. Friedman (1995) is one study that has investigated this relationship. She finds that greater industrial diversity is positively associated with the presence of high-growth small

start-ups, but she measures diversity as the standard deviation of the percentages employed within each industry rather than using the Herfindahl index as in this paper.[10]

RPICH, real per capita income growth, is included to account for the change in wealth and income in each state. It is calculated as the log difference in real per capita income from 1999 to 2000, and it is based on income data from REIS. RPICH is expected to be positively associated with new firm formation for two reasons. First, states with higher growth in disposable income have more income available to be spent on the output from new firms, thus increasing demand in a way similar to population growth. Second, the entrepreneurs themselves have a greater amount of income to spend on starting and financing their new businesses.

The fourth category in Table 3.2, owner-occupied housing, represents an earlier trend in this line of research of using the percentage of owner-occupied housing as a determinant of new firm formation. This idea was based on the assumption that the key method of funding a new business is for the entrepreneur to use his or her home as collateral. Essentially, this category is measuring the availability or access to financing. Instead of using owner-occupied housing, this model relies on the total dollar amount of venture capital financing provided in each state (FIN). These data come from the PricewaterhouseCoopers Moneytree Survey, which began collecting data on a quarterly basis in 1995. It is the only industry-endorsed research effort on venture capital investment activity in the U.S. The a priori expectation for this variable is positive since

[10] Another interesting note on Freidman's study is that she tests for an inverse U-shaped relationship between diversity and high-growth small firms, speculating that excessive diversity my hinder new firms if no industry provides critical mass. Her results, however, do not support this hypothesis.

a greater availability of venture capital financing should encourage a higher rate of new business creation. Specifically, FIN is the average dollar amount of venture capital provided in each state over the three-year period prior to 2001.[11] Focusing on venture capital in particular, as opposed to including local funding sources, captures an additional aspect of working with venture capital firms. That is, venture capitalists tend to provide management advice and information sources to the entrepreneurs they finance, an advantage that local banks do not typically provide.

EDUC is included to represent the skill and education level of the labor force in each state. It is measured as the percentage of the state population that has at least a bachelor's degree and these figures come from the U.S. Statistical Abstracts. EDUC is expected to be positively related to new firm formation since it is generally thought that a state with a higher education level fosters more entrepreneurial growth.

To address the argument that recessions and high unemployment "push" more people into starting their own businesses as an alternative to having no job at all, the state unemployment rate, U, is added as an explanatory variable. Unemployment varies considerably across states and regions so it is important to capture its influence on new firm formation. These data come from the Bureau of Labor Statistics Web site (see Table A3 in the Appendix for the full Web site address). The expectation is that a high unemployment rate is associated with a high rate of new firm formation, but it should

[11] Since there was a large amount of variability from year to year in the amount of financing provided, it was decided that using the three-year average of financing would be more appropriate than simply the one-year lagged value as with the other variables. A couple of states received no financing during this three-year period, but they were assigned a value of one dollar. Doing this allows these few observations to be kept in the analysis once the variable is changed to natural log form.

also be recognized that low unemployment may reduce the risks for entrepreneurs if they know they could find another job easily should their new business fail. It is possible that the prosperity associated with an expansion and low unemployment may "pull" more entrepreneurs into starting a new business to take advantage of market opportunities. Since unemployment is likely to be related to industrial diversity (Izraeli and Murphy, 2003, find that greater diversity is associated with lower unemployment), potential interaction terms are investigated in the econometrics section.

A measure of environmental contamination in each state, ENV, was added to the model and indirectly reflects the multi-faceted role government plays in promoting new business. The inclusion of an environmental contamination variable is an innovation relative to previous literature, so its resulting influence is especially interesting. ENV is the log of the number of hazardous waste sites located in each state in 2000 divided by the 2000 population.[12] These hazardous waste sites are those included on the National Priority List of the Federal Superfund program and the listing of sites per state is provided by the Statistical Abstract of the United States. While hazardous waste sites mostly influence the actual location decision, they can also influence the business start-up decision in a couple of ways. For one, an entrepreneur likely thinks of the long-term future and is less likely to locate their new business, and therefore their family, near hazardous waste sites. Amenities, or disamenities in this case of hazardous waste sites, also influence a new firm's ability to attract and retain employees, a key issue for a new

[12] The states with the highest number of hazardous waste sites were also those with the largest populations, so the number of sites was divided by the population to control for this.

start-up's early success (Friedman 1995). In addition, entrepreneurs may view a high number of hazardous waste sites as a potential spillover cost to them in terms of state taxes. Cebula (2005) finds that the number of hazardous waste sites has a highly significant and negative impact on state in-migration rates, further motivating the inclusion of this variable in the model for this paper. It has been documented that entrepreneurs do not commonly migrate in order to start their new businesses (see Reynolds 1988, for example), but they are affected by the migration of potential employees, as Friedman points out, and therefore the presence of hazardous waste sites may influence start-up decisions. A measure of such environmental damage has not been included in any previous literature on new firm formation.

As mentioned, state governments play a multi-faceted role in promoting business and one of the most difficult concepts to measure and incorporate into this type of analysis is the overall state policy environment for new business start-ups. Yet this aspect is very important, as certain policy initiatives may be the deciding factor for many entrepreneurs when choosing whether or not to begin a new venture. It would be very time-consuming to research and compile the various policies for each state, so a published index is used to measure the business climate for new firms. This variable, CLIM, is a ranking from 1 to 50 of the public policy environment in each state with 1 being the most positive toward small business start-ups and 50 being the most hostile or restrictive. This index is compiled annually by the Small Business and Entrepreneurship Council and published on their Web site (see Table A3 for the complete Web address). It

is based on seventeen major government-related or government-imposed costs:

- personal income tax
- corporate income tax
- sales tax
- unemployment tax
- electricity costs
- crime rates
- number of bureaucrats
- Internet tax
- state minimum wage
- capital gains tax
- property tax
- death tax
- health insurance tax
- workers' compensation costs
- right to work status
- tax limitation status
- gas tax

Table A4 in the Appendix lists the rankings for 2000. CLIM is expected to be negatively associated with the dependent variable, that is a ranking low in number will be associated with a higher rate of new firm creation.

Finally, WEB is included in the model both as a literal measure of Internet access in each state and as a proxy for the availability and application of new technology in general. It is specifically measured as the percentage of zip codes in each state that had at least one provider of high-speed Internet access in 2000. The Federal Communication Commission began reporting this statistic in 1999. This variable has not been analyzed in much of the previous literature on firm formation, but it is of increasing importance to small businesses as e-commerce continues its rapid growth. High-speed Internet access greatly increases the markets available for both purchasing inputs and selling outputs. It is also seen as a way to share and gain industry knowledge and expertise, thus allowing

for agglomeration effects despite a lack of geographical proximity. Therefore, a higher value of WEB should be positively associated with a higher firm birth rate.

Table 3.3 – Descriptive Statistics

Variable	Mean	Std. Deviation	Minimum	Maximum
NEW_SM	11,418.67	13,152.42	1,453.00	72,796.00
NEW_MD	654.80	781.17	63.00	4,662.00
CLIM	25.59	14.72	1.00	50.00
DIV	947.81	120.60	785.86	1390.57
EDUC	24.98	4.34	15.30	34.60
FIN	351,000,000	960,000,000	1.00	6,460,000,000
MES	16.88	2.33	10.62	21.23
POPCH	0.01	0.01	0.00	0.04
RPICH	0.03	0.01	0.00	0.07
U	3.91	0.95	2.20	6.69
WEB	74.82	17.63	22.00	100.00
ENV	26.08	26.20	1.00	113.00
N	*49*			

Descriptive statistics and correlations for these variables are provided in Tables 3.3 and 3.4, respectively. More detailed descriptions of the data sources are listed in Table A3 of the Appendix. In Table 3.3, the statistics for the following variables are reported as levels but in the actual analyses, the log form of each is used: NEW_SM, NEW_MD, FIN, and ENV.

From the simple correlation between the independent variables and the dependent variable, it appears that all of the potential determinants have the expected or logical sign on them. There are no correlation values that really stand out as problematic, but there are a few that are relatively higher than the rest. The potential collinearity between venture capital financing (FIN) and mean existing establishment size (MES) is one that is addressed in the next section.[13]

[13] There were two other relatively high correlations that did not seem to pose problems in the model. DIV and EDUC had a correlation of –0.608 but both turned out to be statistically significant. EDUC and RPICH had a correlation of –0.605 but when one or the other was omitted from the model, their statistical significance did not change.

Table 3.4 - Correlations

	NEW_SM	NEW_MD	CLIM	DIV	EDUC	FIN	MES	POPCH	RPICH	U	WEB	ENV
NEW_SM	1.000	-										
NEW_MD	-	1.000										
CLIM	-0.069	-0.099	1.000									
DIV	-0.394	-0.328	-0.155	1.000								
EDUC	0.188	0.102	0.207	-0.608	1.000							
FIN	0.721	0.712	0.132	-0.373	0.413	1.000						
MES	0.565	0.649	-0.133	0.070	0.007	0.640	1.000					
POPCH	0.225	0.201	-0.267	0.077	0.052	0.270	0.315	1.000				
RPICH	0.215	0.144	0.028	-0.538	0.605	0.270	-0.058	0.115	1.000			
U	-0.043	-0.041	0.034	0.015	-0.371	-0.117	-0.251	-0.155	-0.319	1.000		
WEB	0.369	0.310	0.034	-0.237	0.331	0.419	0.294	0.270	0.439	-0.295	1.000	
ENV	-0.323	-0.409	0.312	-0.253	0.265	-0.050	-0.417	-0.380	0.257	-0.047	0.109	1.000

4. Econometric Issues and Results

Consistent with previous literature, ordinary least squares regression analysis is used to model the results for each of the two size categories. As noted in the variable descriptions, each independent variable is lagged one year to reflect the reality that business decisions take time and are likely based on the conditions present in the year prior to the actual start-up of the new business. The sample size for all of the models is 49 states due to the fact that North Dakota had zero hazardous waste sites in 2000, and thus the log of ENV yields a missing value for that state.[14]

The results are presented in Tables 3.5 and 3.6 for small- and medium-sized new business establishments respectively. Again, small-sized businesses are those with fewer than 20 employees and medium-sized businesses are those with 20 to 99 employees. Model 1 represents the initial analysis while models 2 and 3 are estimated to investigate the potential problem of collinearity between FIN and MES. The correlation matrix shows that FIN and MES have the highest correlation values relative to the dependent variable, yet MES does not turn out to be significant in Model 1. In addition, when MES is omitted, FIN retains its statistical significance (Model 2), but when FIN is removed, MES becomes significant (Model 3).

To address this issue, a simple two-variable regression is estimated using FIN as the dependent variable and MES as the explanatory variable. It is logical to expect that states with larger businesses on average are more likely to attract the attention of venture

[14] Regressions were estimated with an arbitrarily assigned value of 1 hazardous waste site for North Dakota. The results were nearly identical to the results without that state included, so it was decided that omitting North Dakota would not affect the results.

capital investors. The residual from this analysis captures the variation in FIN that is unexplained by MES. So RESID_FIN is substituted for FIN in the original model, and then both MES and RESID_FIN turn out to be statistically significant (Model 4).

This model assumes that the constant is the same for all states, which may not always be true. However, when the model is estimated using dummy variables for states that appear to have relatively larger residuals, no sizable changes are seen in the results. Therefore, no state dummy variables are left in the model.[15]

All models except Model 3 appeared to have heteroskedasticity issues based on the White residual test, so the models are estimated using the White correction for heteroskedasticity in Eviews. The interpretation of the results focuses on Model 4 from both of the two size categories. In both tables of results, the t-statistics are reported in parentheses below the coefficients.

[15] Interaction terms were also considered when estimating this model. Nearly all interactions yielded insignificant coefficients. Those that did turn out to be significant did not have clearly justified relationships between them. Therefore, no interaction terms were added to the final models.

Table 3.5 – Regression Results: New Small Firms as Dependent Variable
(Fewer Than 20 Employees)

Independent Variable	Model 1	Model 2	Model 3	Model 4
CLIM	-0.005 (-0.809)	0.000 (0.037)	-0.005 (-0.847)	-0.005 (-0.809)
DIV	-0.003** (-2.645)	-0.004*** (-3.243)	-0.003*** (-2.759)	-0.003** (-2.645)
EDUC	-0.056** (-2.088)	-0.020 (-0.594)	-0.056** (-2.192)	-0.056** (-2.088)
FIN	0.174** (2.405)		0.178*** (3.380)	
RESID_FIN				0.174** (2.405)
MES	0.008 (0.109)	0.192*** (3.734)		0.200*** (3.543)
POPCH	-14.053 (-1.114)	-2.914 (-0.196)	-14.223 (-1.183)	-14.053 (-1.114)
RPICH	3.268 (0.378)	4.945 (0.480)	3.092 (0.366)	3.268 (0.378)
U	-0.029 (-0.321)	0.110 (0.890)	-0.034 (-0.401)	-0.029 (-0.321)
WEB	0.008 (1.278)	0.010 (1.461)	0.008 (1.285)	0.008 (1.278)
ENV	-0.473** (-2.258)	-0.353* (-1.997)	-0.481** (-2.674)	-0.473** (-2.258)
C	3.609 (1.182)	3.994 (1.378)	3.606 (1.207)	3.335 (1.079)
N = 49				
Adjusted R-squared	*0.64*	*0.50*	*0.65*	*0.64*
F-statistic	*9.42*	*6.29*	*10.73*	*9.42*

*** *indicates significance at the 1% level*
** *indicates significance at the 5% level*
* *indicates significance at the 10% level*

Table 3.6 – Regression Results: New Medium Firms as Dependent Variable
(20 to 99 Employees)

Independent Variable	Model 1	Model 2	Model 3	Model 4
CLIM	-0.005 (-0.897)	0.000 (0.027)	-0.006 (-1.050)	-0.005 (-0.897)
DIV	-0.003*** (-3.005)	-0.004*** (-3.506)	-0.003*** (-2.988)	-0.003*** (-3.005)
EDUC	-0.072*** (-3.123)	-0.036 (-1.103)	-0.078*** (-3.441)	-0.072*** (-3.123)
FIN	0.176** (2.659)		0.205*** (3.955)	
RESID_FIN				0.176*** (2.659)
MES	0.058 (0.879)	0.244*** (4.904)		0.253*** (4.865)
POPCH	-22.143* (-1.943)	-10.837 (-0.750)	-23.421** (-2.188)	-22.143* (-1.943)
RPICH	4.549 (0.522)	6.251 (0.624)	3.223 (0.381)	4.549 (0.522)
U	-0.047 (-0.524)	0.094 (0.784)	-0.085 (-1.036)	-0.047 (-0.524)
WEB	0.005 (0.936)	0.008 (1.137)	0.006 (1.059)	0.005 (0.936)
ENV	-0.569*** (-2.756)	-0.447** (-2.603)	-0.625*** (-3.696)	-0.569*** (-2.756)
C	-0.464 (-0.154)	-0.074 (-0.026)	-0.484 (-0.170)	-0.743 (-0.243)
N = 49				
Adjusted R-squared	*0.71*	*0.58*	*0.71*	*0.71*
F-statistic	*12.76*	*8.48*	*14.11*	*12.76*

*** *indicates significance at the 1% level*
** *indicates significance at the 5% level*
* *indicates significance at the 10% level*

The F-statistic for both models is statistically significant at the one percent level indicating that the independent variables explain the variation in new small firm establishments reasonably well. The adjusted R-squared values of 63.7 and 71.0 mean that approximately 60 to 70 percent of the variation in new small firm establishments is explained by the determinants included in the model. The percent of explained variation in the previously mentioned literature ranges widely from 49 to 86 percent, with the majority falling in the 60 percent range. Thus, the adjusted R-squared values found here are very similar.

As expected, the number of hazardous waste sites (ENV) has a statistically significant negative impact on new firm formation, indicating that entrepreneurs are likely to avoid starting a business in locations with a greater number of hazardous waste sites. A 10 percent increase in the number of hazardous waste sites per person is associated with a 4.7 percent decrease in small firm formation. For medium-sized start-ups, a 10 percent increase in hazardous waste sites corresponds to a 5.7 percent decrease in firm formation.

The statistically significant positive coefficient on the venture capital finance variable (FIN) is also in line with expectations. One would expect that a higher dollar amount of financing would signal a greater availability of capital and therefore motivate more entrepreneurs to start businesses. Based on the coefficient for FIN, a 10 percent increase in the dollar amount of venture capital financing provided is associated with a nearly two percent increase in both small and medium firm formation.

The significant negative coefficient for education level (EDUC) appears to be counterintuitive at first, but makes more sense when considering that certain industries, especially manufacturers, rely on a large, less-educated workforce (Lee et al. 2004). The a priori expected sign for education level was positive since it is thought that a college education provides an entrepreneur with knowledge and expertise to form a new business. However, the actual resulting sign on this coefficient is negative, indicating the presence of lower-educated, and therefore cheaper, labor inputs has more influence on the start-up decision. Reynolds et al. (1995) also finds that higher education levels are negatively associated with new firm formation.

The a priori expectation for industrial diversity (DIV) was ambiguous. Reynolds et al. (1995) finds higher firm formation rates to be associated with industry specialization, but Friedman (1995) finds new firms to be associated with diversity rather than specialization. In this model, DIV turns out to be a statistically significant negative determinant of new firm formation. Recall that a higher Herfindahl index value indicates industrial specialization, so industrial specialization appears to be associated with lower firm formation rates. Greater diversity promotes new small firm formation and new medium firm formation according to the results of these models and this is consistent with the findings of Friedman (1995). New small firms probably find it easier to compete when they can exploit niche markets and serve a wide variety of industries. In an industrially concentrated state, a new firm may be unable to either compete with a large number of existing businesses in the same industry or to overcome barriers to entry.

As with industrial diversity, the a priori expectation for mean existing establishment size (MES) was indeterminate. This estimation indicates that MES has a statistically significant positive impact on new firm formation after adjusting the model for the multicollinearity between FIN and MES. This finding is consistent with Sutaria and Hicks (2004) and as they interpret it, this means that new firms benefit in some way from the presence of larger firms in the area. The large firms may purchase inputs from and outsource work to small neighboring firms, or large firms may share contacts and financial support with the entrepreneurs if they do not view the new start-ups as competitors. Even if the large and small firms are competitors, Friedman (1995) points out that small firms making the same product can help meet peak demand in highly fluctuating markets. In general, the presence of large firms likely offers a more stable environment in which new firms can prosper.

The above results describe the models for both size categories examined. The only difference in significant determinants between the two models is that population change (POPCH) is statistically significant and negative in the model of medium firm formation rates. Although this result is not common in the literature, only 4.3 percent of new firm births fall into the medium category so it is possible that migrants to a state are moving there either to take a job with an existing business or they are starting a smaller business. Other than this variable, there is very little difference between the determinants of small and medium firm formation rates so the separation by initial firm size does not yield any insights as hoped. Since the vast majority of new firms are in the small category, more emphasis should be placed on those results.

It is interesting to examine why some of the variables actually turn out to be insignificant. Growth in real personal per capita income (RPICH) is insignificant and Armington and Acs (2002) suggest that this is because personal wealth is relatively less important today in founding new firms. Sutaria and Hicks (2004) further rationalize that supply chains, especially in the manufacturing industry, and markets in general are increasingly global and therefore it is understandable why local income is less important to new firms.

Unemployment (U) is also insignificant and this is most likely due to countervailing influences, which are difficult to isolate, canceling each other out. While high unemployment may push some entrepreneurs to start new ventures due to the ease of hiring labor, it may deter others who view it as a greater risk and a cause of lower demand for output. The insignificance of the business climate variable (CLIM) may also be a result of opposite influences. A state with a very friendly business climate may also have a very highly competitive environment that would be difficult for new firms to compete in. It is also possible that the measure of business climate used was too highly aggregated and a more narrow measure of the state tax structures may have yielded better results.

Finally, it is somewhat unexpected to see that high-speed Internet access (WEB) does not have a significant impact on new firm formation rates given all of the talk about the New Economy and how important technology and access to information are in business today. It is not easily explained why this measure of Internet access fails to show an influence on business start-ups. Perhaps a different definition of Internet access

other than high-speed only would have yielded different results. Separate regressions for different industries may have provided more insight as well since certain service industries may rely heavily on Internet access while other manufacturing industries may not be affected by a lack of high-speed Internet access.

5. Conclusions and Policy Implications

The goal of this chapter was to learn which regional economic and socioeconomic characteristics most strongly influence entrepreneurs to start a new business. Based on a model of the firm formation rates of 49 states in 2001, it appears that five characteristics emerge as significant determinants: industrial diversity, education level, availability of financing, average size of existing establishments, and presence of environmental contamination. Consistent with Friedman (1995), greater industrial diversity is associated with higher rates of small firm formation. Industrial diversity is not commonly included in models of new firm formation rates, so this research builds on Friedman's by using the Herfindahl index and further strengthens the relatively newer theory that diversity rather than specialization promotes entrepreneurship.

Another important contribution from this research effort is the finding that environmental contamination is associated with lower rates of small firm formation. This is a new finding as no previous study appears to have investigated the relationship between the prevalence of hazardous waste sites and small business start-ups. The findings for finance and for mean existing establishment size are consistent with Sutaria and Hicks (2004) and the finding for education level is consistent with Reynolds, et al.

(1995). The relationship between each of these factors and the rate of new firm formation is likely much more complex than a simple positive or negative coefficient can reflect, but nonetheless, some valuable policy implications can be taken away from this research.

The positive relationship between industrial diversity and small firm formation suggests that a state or region should be open to and encouraging of the development of all industries, not just their primary industry of expertise. This may seem to contradict the large literature on spatial agglomeration and spillover benefits among like industries, but it could be due to the larger geographic unit of observation. That is, while firms in the same industry and same city may benefit from agglomeration and this contributes to growth, firms within the larger regional context may grow and prosper better when there is greater diversity and a wider variety of industries with which to do business. Basically, this relationship should be a caution to policymakers not to put all of their eggs in one basket. There could be a tendency for policymakers to believe that if their state has one particularly profitable industry, then they should focus their efforts and resources on continuing to develop new firms within that industry. This research, however, suggests that such industrial concentration may actually hinder the start-up of small firms.

The negative relationship between education and new firm formation is also complex and certainly does not mean that states should stop encouraging higher education levels. One way to translate this result into policy could be to promote entrepreneurship at an earlier age and at lower education levels. For example, educational seminars on business formation for high school students may be more

beneficial for overall growth than such seminars at the college level. In addition, the process of awarding state aid and grants to new businesses should not discriminate against those entrepreneurs without college degrees. Literature from other disciplines on the personality characteristics of entrepreneurs would tie in well with this discussion. This finding of a negative coefficient on education could be interpreted to mean that the lack of a college degree does not necessarily constrain an entrepreneur if they have the needed expertise, ambition, and resources to start their own business. Bill Gates is a famous example of a successful entrepreneur who did not attain a college degree.

The positive coefficient on the availability of venture capital financing is very logical and expected, and this result is the simplest to translate into policy. Entrepreneurs are more likely to start up new businesses when they believe capital is more readily available. States can contribute their own resources, streamline application processes, or offer incentives to venture capital providers to increase the availability of financing.

While mean existing establishment size is positively associated with new firm formation, it does not mean that a state must have a high number of very large businesses. Rather it indicates that the potentially beneficial relationships between the existing large businesses and the new start-ups should be fostered. Policymakers could increase awareness among larger businesses of the potential benefits of their interactions with small start-ups. For example, one such benefit may be that small start-ups are lower-cost suppliers of intermediate inputs (Sutaria and Hicks, 2004). A mentoring program of sorts could encourage entrepreneurship and the larger businesses themselves would eventually benefit from the overall growth in their state's economy.

Finally, the negative relationship found between hazardous waste sites and new firm formation should further motivate policymakers to minimize the prevalence of environmental risks in their states. Besides the obvious health and safety reasons, a cleaner environment is a signal of long-term viability for businesses as well as individuals. Cleaning up polluted neighborhoods and revitalizing old contaminated industrial areas provides new commercial locales for business start-ups. This trend is already popular and should be continued according to the results of this research.

As discussed at the beginning of this chapter, entrepreneurs contribute to regional economic growth in various important ways and therefore policies to encourage new firm formation are warranted. However, inward investment from outside established businesses should not be completely ignored in order to foster entrepreneurial growth. New branch plants from existing businesses can expose local firms to state-of-the-art technology and management methods, benefiting all local firms and contributing to growth through this "demonstration effect" (Armstrong and Taylor 2000, p. 264).

Furthermore, regional policy to promote firm *survival* rates is just as important as promoting firm *formation* rates in order to address the common criticisms of small business start-ups. New firms tend to provide less stable employment in general, so policies aimed at helping entrepreneurs sustain their new business ventures through the first three years and beyond can contribute to more stable employment for the overall region. It is likely that many of the determinants of regional firm formation rates would also have significant influences on survival rates, but the concept is different enough to

deserve a separate and distinct econometric model. Thus, this is one suggestion for further research that arises from this research effort.

The possibilities for future research on this topic are numerous and varied. One obvious direction for further research is to study panel data over time. The main limitation of this study is that it uses 2001 data only, for reasons mentioned earlier. This makes it difficult to compare it to studies that span different decades. To further make this study more comparable to other studies, additional variables could be included, such as the proportion of immigrants. Some studies have found immigration rates to be significantly associated with new firm formation rates.

Another appealing topic for future research would be an analysis of the cost-effectiveness of the various suggested policies for promoting new business start-ups. Entrepreneurship is a very popular and trendy subject these days and it would interesting to see the actual effects in terms of state employment figures and GDP growth from the resources being devoted to fostering small businesses. However, just as entrepreneurship itself is difficult to measure, the results of such policies may be just as difficult to quantify. If entrepreneurs contribute more through informal economic activities or in the Schumpeterian sense of encouraging innovation and competition, then these benefits may not be directly reflected in our standard economic measures of employment and GDP.

IV. Conclusion

Industrial diversity is the underlying theme of this thesis that examines two important issues in regional economics. The first goal of this thesis was to examine state-level recession duration and its determinants. Nonparametric duration analysis shows that the recessions between 1979:Q1 and 1997:Q4 exhibit positive duration dependence. This indicates that the longer a recession persists, the more likely it is to end rather than continue. Full parametric analysis then shows that numerous factors are significantly related to longer recessions. The factors associated with longer recessions include decreases in industrial diversity, high unemployment, decreases in real income per capita, high proportions of non-white workers, small total populations, and increases in population growth.

Using some of the same independent variables, the second goal of this thesis was to investigate the determinants of new firm formation at the state level. Based on a model of the small firm formation rates of 49 states in 2001, five characteristics are significantly associated with higher small firm formation rates: greater industrial diversity, lower education levels, greater availability of financing, larger average size of existing establishments, and fewer environmental hazardous waste sites. This last finding in particular regarding the influence of hazardous waste sites is notable since no previous studies, to the author's knowledge, have included such an environmental contamination measure as a potential determinant.

Why are some state economies able to outperform those of other states? This question was posed in the introduction and from this research effort, it is clear that industrial diversity is one potential answer. Industrial diversity does seem to play an important role in the functioning of state economies. It appears that increases in industrial diversity are associated with shorter recessions, and higher levels of diversity are associated with higher rates of new small firm formation. Therefore, state policies aimed at increasing industrial diversity are justifiable and would likely have a positive influence on the growth and performance of state economies.

Appendix

Table A1 – Data Descriptions and Sources for Chapter 2

Variable	Description	Source
DURATION	Number of quarters over which a state's recession persisted.	Crone (2002)
DIV	Industrial diversity in a state in the year prior to the recession. Calculated using the Herfindahl Index.	County Business Patterns, www.census.gov/epcd/cbp/view/cbpview.html
DIVCH	Change in industrial diversity. Calculated as the change in the Herfindahl index from year prior to recession to year of recession.	County Business Patterns, www.census.gov/epcd/cbp/view/cbpview.html
U	State unemployment rate in the year prior to the recession.	Bureau of Labor Statistics, www.bls.gov/
RPICH	Change in state real per capita personal income. Change calculated as log difference in real income from year prior to recession to year of recession. Real income calculated from nominal income using annual CPI based on All Urban Consumers.	Regional Economic Information System, www.bea.doc.gov/bea/regional/reis/ (income) Bureau of Labor Statistics, http://www.bls.gov/cpi/home.htm (CPI)
NWT	Percentage of a state's working-age population (15-64 years) that is non-white. Calculated from Census data.	U.S. Census Bureau, www.census.gov/popest/archives
TEEN	Percentage of a state's working-age population that is 15-19 years of age. Calculated from Census data.	U.S. Census Bureau, www.census.gov/popest/archives
OVER65	Percentage of a state's total population that is 65 years or older. Calculated from Census data.	U.S. Census Bureau, www.census.gov/popest/archives

Table A1 – Data Descriptions and Sources for Chapter 2 (continued)

Variable	Description	Source
POP	Log of total state population in the year prior to the recession.	Regional Economic Information System, www.bea.doc.gov/bea/regional/reis/
POPCH	Rate of population growth in a state. Change calculated as log difference in annual population from year prior to recession to year of recession.	Regional Economic Information System, www.bea.doc.gov/bea/regional/reis/
DENS	Population density. Calculated as persons per square mile.	Regional Economic Information System, www.bea.doc.gov/bea/regional/reis/

Table A2 – Recession Duration Data and Nonparametric Hazard, Survivor, and Integrated Hazard Estimates

Duration in Quarters t_j	Completed Recessions j_j	Uncompleted Recessions n_j	Hazard Estimate $\hat{\lambda}(t_j)$	Survivor Estimate $\hat{S}(t_j)$	Integrated Hazard $\Lambda(t)$
2	20	148	0.135	0.865	0.145
3	17	128	0.133	0.750	0.288
4	14	111	0.126	0.655	0.423
5	25	97	0.258	0.486	0.721
6	18	72	0.250	0.365	1.008
7	10	54	0.185	0.297	1.213
8	9	44	0.205	0.236	1.442
9	5	35	0.143	0.203	1.596
10	2	30	0.067	0.189	1.665
11	4	28	0.143	0.162	1.819
12	2	24	0.083	0.149	1.906
13	3	22	0.136	0.128	2.053
14	6	19	0.316	0.088	2.432
15	5	13	0.385	0.054	2.918
16	2	8	0.250	0.041	3.205
17	1	6	0.167	0.034	3.388
18	1	5	0.200	0.027	3.611
19	0	4	0.000	0.027	3.611
20	0	4	0.000	0.027	3.611
21	1	4	0.250	0.020	3.899
30	1	3	0.333	0.014	4.304
52	1	2	0.500	0.007	4.997

Table A3 – Data Descriptions and Sources for Chapter 3

Variable	Description	Source
NEW_SM NEW_MD	Annual new firm establishments (log form). SM (small) = 1-19 employees MD (medium) = 20-99 employees	Statistics of U.S. Businesses, www.census.gov/csd/susb/susb.htm
CLIM	Annual rankings of state policy climate for entrepreneurship (lag=1). Index based on 17 major government-related costs: personal income tax, capital gains tax, corporate income tax, property tax, sales tax, death tax, unemployment tax, health insurance tax, electricity costs, workers' compensation costs, crime rates, right to work status, number of bureaucrats, tax limitation status, Internet tax, gas tax, and state minimum wage.	Small Business and Entrepreneurship Council, www.sbsc.org/LatestNews_Action.asp?FormMode=Releases&ID=195
DIV	Industrial diversity in a state (lag=1). Measured as the Herfindahl Index.	County Business Patterns, www.census.gov/epcd/cbp/view/cbpview.html
EDUC	Percentage of state population with at least a bachelor's degree (lag=1).	U.S. Census Bureau, Statistical Abstract of the United States.
FIN	Dollar amount of venture capital financing provided in a state (average of lag=1, lag=2, and lag=3; log form). Sum of venture capital provided during "seed" stage (<18 months) and "early" stage (< 3 years).	PricewaterhouseCoopers Moneytree Survey, http://www.pwcmoneytree.com/moneytree/index.jsp
MES	Mean establishment size (lag=1). Calculated as employment divided by total establishments.	Statistics of U.S. Businesses, www.census.gov/csd/susb/susb.htm
POPCH	Rate of population growth in a state (lag=1). Change calculated as log difference in annual population.	Regional Economic Information System, www.bea.doc.gov/bea/regional/reis/

Table A3 – Data Descriptions and Sources for Chapter 3 (continued)

Variable	Description	Source
RPICH	Change in state real per capita personal income (lag=1). Change calculated as log difference in real income. Real income calculated from nominal income using annual CPI based on All Urban Consumers.	Regional Economic Information System, www.bea.doc.gov/bea/regional/reis/ (income) Bureau of Labor Statistics, www.bls.gov/cpi/home.htm (CPI)
U	State unemployment rate (lag=1).	Bureau of Labor Statistics, www.bls.gov/sae/home.htm
WEB	Percentage of zip codes in a state with at least one provider of high-speed Internet access (lag=1).	Federal Communication Commission, www.fcc.gov/web/iatd/comp.html
ENV	Number of hazardous waste sites on National Priority List in each state (lag=1). Calculated as number of sites divided by state population (log form).	U.S. Census Bureau, Statistical Abstract of the United States.

Table A4 – 2000 Small Business Survival Index Rankings

Rank	State	Rank	State
1	South Dakota	26	Delaware
2	Nevada	27	Massachusetts
3	Wyoming	28	West Virginia
4	New Hampshire	29	Wisconsin
5	Texas	30	Nebraska
6	Florida	31	Oklahoma
7	Washington	32	Idaho
8	Alabama	33	Utah
9	Michigan	34	Arkansas
10	Mississippi	35	Connecticut
11	Tennessee	36	Kansas
12	Alaska	37	Vermont
13	Indiana	38	Iowa
14	Missouri	39	California
15	South Carolina	40	New York
16	Colorado	41	North Carolina
17	Virginia	42	Maine
18	Louisiana	43	New Jersey
19	Illinois	44	Oregon
20	North Dakota	45	Montana
21	Georgia	46	Ohio
22	Maryland	47	Minnesota
23	Arizona	48	New Mexico
24	Pennsylvania	49	Rhode Island
25	Kentucky	50	Hawaii

Source: Small Business and Entrepreneurship Council.

Bibliography

Armington, Catherine and Zoltan Acs. "The Determinants of Regional Variation in New Firm Formation." *Regional Studies*, Vol. 36, No. 1 (February 2002): pp. 33-45.

Armstrong, Harvey and Jim Taylor. *Regional Economics and Policy*, Third Edition. Malden, MA: Blackwell Publishers, 2000.

Audretsch, David B. and Max Keilbach. "Entrepreneurship Capital and Economic Performance." *Regional Studies*, Vol. 38, No. 8 (November 2004): pp. 949-959.

Carlino, Gerald and Keith Sill. "Regional Income Fluctuations: Common Trends and Common Cycles." *Review of Economics and Statistics*, Vol. 83, No. 3 (2001): pp. 446-456.

Cebula, Richard. "Internal Migration Determinants in the U.S., 1999-2002." Armstrong Atlantic State University Working Paper (2005).

Crone, Theodore M. "Coincident Economic Indexes for the 50 States." Federal Reserve Bank of Philadelphia Working Paper No. 02-7 (2002).

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. *Job Creation and Destruction*. Cambridge: The MIT Press, 1996.

Friedman, Judith. "The Effects of Industrial Structure and Resources Upon the Distribution of Fast-growing Small Firms Among U.S. Urbanized Areas." *Urban Studies*, Vol. 32, No. 6 (1995): pp. 863-883.

Garcia-Mila, Teresa and Therese J. McGuire. "Industrial Mix as a Factor in the Growth and Variability of States' Economies." *Regional Science and Urban Economics*, Vol. 23 (1993): pp. 731-748.

Georgellis, Yannis and Howard J. Wall. "What Makes a Region Entrepreneurial? Evidence from Britain." Federal Reserve Bank of St. Louis Working Paper (June 1999).

Gilbert, Brett Anitra, David B. Audretsch, and Patricia P. McDougall. "The Emergence of Entrepreneurship Policy." *Small Business Economics*, Vol. 22 (2004): pp. 313-323.

Greene, William. *Econometric Analysis*, New York: MacMillan, 1993.

Henderson, Jason. "Building the Rural Economy with High-Growth Entrepreneurs." Federal Reserve Bank of Kansas City, *Economic Review* (Third Quarter 2002): pp. 45-70.

Izraeli, Oded and Kevin J. Murphy. "The Effect of Industrial Diversity on State Unemployment Rate and Per Capita Income." *The Annals of Regional Science*, Vol. 37 (2003): pp. 1-14.

Kiefer, Nicholas M. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature*, Vol. 26 (June 1988): pp. 646-679.

Lee, Sam Youl, Richard Florida, and Zoltan J. Acs. "Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation." *Regional Studies*, Vol. 38, No. 8 (November 2004): pp. 879-891.

Malecki, Edward J. "Entrepreneurship in Regional and Local Development." *International Regional Science Review*, Vol. 16, Nos. 1 & 2 (1994): pp. 119-153.

Owyang, Michael T., Jeremy Piger, and Howard J. Wall. "Business Cycle Phases in U.S. States." Federal Reserve Bank of St. Louis Working Paper (2004).

Parker, Simon C. *The Economics of Self-Employment and Entrepreneurship*, Cambridge, UK: Cambridge University Press, 2004.

Reynolds, Paul. "Organizational Births: Perspectives on the Emergence of New Firms." *Academy of Management Best Papers Proceedings* (1988): pp. 69-73.

Reynolds, Paul, David J. Storey, and Paul Westhead. "Cross-National Comparisons of the Variation in New Firm Formation Rates." *Regional Studies*, Vol. 28, No. 4 (1995): pp. 443-456.

Simon, Curtis J. and Clark Nardinelli. "Does Industrial Diversity Always Reduce Unemployment? Evidence from the Great Depression and After." *Economic Inquiry*, Vol. 30, No. 2 (April 1992): pp. 384-397.

Stock, James H. and Mark W. Watson. "New Indexes of Coincident and Leading Economic Indicators." *NBER Macroeconomics Annual* (1989): pp. 351-394.

Storey, D.J. *Understanding the Small Business Sector.* London: Routledge, 1994.

Sutaria, Vinod and Donald A. Hicks. "New Firm Formation: Dynamics and Determinants." *Annals of Regional Science*, Vol. 38 (2004): pp. 241-262.

Wagner, Gary A. and Erick M. Elder. "Recessions and Rainy Day Funds in U.S. States: What Are the Odds Your State is Saving Enough?" Duquesne University Working Paper (December 2004).

Wall, Howard J. "Entrepreneurship and the Deregulation of Banking." *Economics Letters*, Vol. 82 (2004): pp. 333-339.

Wall, Howard J. and Gylfi Zoega. "U.S. Regional Business Cycles and the Natural Rate of Unemployment." Federal Reserve Bank of St. Louis *Review*, Vol. 86, No. 1 (2004): pp. 23-31.

Wundt, Bruce D. "Reevaluating Alternative Measures of Industrial Diversity as Indicators of Regional Cyclical Variations." *The Review of Regional Studies*, Vol. 22, No. 1 (1992): pp. 59-73.