August, 2013

Employment Risk, Returns, and Entrepreneurship

Sarah A. Low
Stephan Weiler, Colorado State University - Fort Collins

Available at: https://works.bepress.com/sarah_low/3/
Employment Risk, Returns, and Entrepreneurship

Abstract

Comparing local employment portfolios against entrepreneurship, this paper finds local wage and salary job market prospects shape incentives for potential entrepreneurs. Entrepreneurship may thus be more attractive in areas featuring high employment risk and/or low returns. This paper contributes to the existing regional employment portfolio literature by using more disaggregated data, at both the county and commuting zone levels. Commuting zones in particular represent a broader spectrum of labor market agglomerations across both rural and urban areas to provide the most stringent and revealing tests of the inter-relationship between local employment portfolios and the choice to pursue entrepreneurship. We find a U-shaped risk/return tradeoff using employment variance and growth, consistent with the literature. We test our hypothesis with a model of regional entrepreneurship, incorporating the employment portfolio variables. This is the first known study to explore the hypothesized relationship between wage and salary employment portfolios and entrepreneurship, effectively synthesizing two previously disparate literatures.
Employment Risk, Returns, and Entrepreneurship

1. Introduction

Research examining the determinants of entrepreneurship often fails to consider the behavior of local labor markets. Entrepreneurship may be more attractive in areas featuring high employment risk and/or low returns. Better understanding how labor market risks and returns incent nascent entrepreneurs will help advance entrepreneurship research and improve the implementation of effective economic development policies and programs. This paper finds, when comparing employment portfolios against regional entrepreneurship, measured empirically with self employment and establishment birth rates, local wage and salary job market prospects indeed shape incentives for potential entrepreneurs.

This paper makes contributions to both the employment portfolio and entrepreneurship literature, as they relate to local labor market research and economic development. The paper is the first known to connect regional employment portfolio theory with entrepreneurship, particularly in terms of the incentive to pursue self employment given the risk/return patterns of local wage and salary positions. Second, this paper represents the first consistently successful application of portfolio theory to labor markets; we consider more localized data than states and metros (commuting zones as well as counties) to evaluate the underlying relationships in the most revealing geographic detail which still represents cohesive labor markets. In particular, commuting zone analysis has the advantage of broadening the context of inquiry to both metropolitan
and non-metropolitan areas through a consistent taxonomy of labor market agglomerations defined by commuting sheds.

In section 2, we set up the tradeoff between paid employment and self employment in counties and commuting zones, drawing on the employment portfolio theory literature, which differs from, but has roots in, financial portfolio analysis. Seminal literature describes a risk/return tradeoff using employment variance and growth and suggests the tradeoff often features an underappreciated U-shape (Chandra, 2002; Lande, 1994; Spelman, 2006). We find evidence of this U-shape for smaller areal units.

Section 3 explores the likely supply-side influence of wage and salary employment portfolios on engagement in entrepreneurial activity by empirically testing whether employment portfolios affect next period’s self employment and establishment birth rates. We test our hypothesis by incorporating the employment portfolio variables into a model of regional entrepreneurship. Findings suggest entrepreneurship is affected by the employment portfolio. By better understanding the determinants of entrepreneurship and how job prospects affect entrepreneurship, policymakers can better foster entrepreneurship. Section 4 summarizes the findings and contributions and concludes with policy recommendations.

2. Portfolio Theory and Regional Applications

2.1. Regional Employment Applications of Portfolio Theory

Regional employment portfolio theory, clearly established in the regional science literature, is the portfolio theoretic approach to economic growth and stability. To this point, the focus of regional employment portfolio analysis has been on finding the
optimal mix of industries in a region (Barry & Kearney, 2006; Barth, Kraft, & Wiest, 1975; Chandra, 2002; Conroy, 1974; Lande, 1994; Spelman, 2006).

Financial portfolio theory proposes that rational investors will optimize their portfolio by diversifying expected risk and return. Regional employment structures can be considered in the same light, where a region’s employment is considered to be its stochastic portfolio of assets (Conroy, 1974; Jackson, 1984). The region affects its employment portfolio, not by buying and selling stocks, but via economic development strategies such as industrial recruitment or business infrastructure provision. The accepted approach in translating financial measures to job markets is to define risk as the variance of employment, while returns are the growth in employment (Lande, 1994).

Financial portfolios have risk and return tradeoffs and feature a strictly increasing risk and return frontier (Markowitz, 1952). Regional employment portfolios, however, feature an isolatable U-shape tradeoff, in which regions on the far left of the U-shape have low or negative employment growth and high employment volatility (perhaps due to a recent plant closure) while those on the far right have higher job growth and high volatility (due to job growth and creative destruction (Schumpeter, 1975[1942]), or high business churn). Research has found many states and metro areas lie in the upper-left of the U-shape (Chandra, 2002; Lande, 1994; Spelman, 2006).

A priori, one might think application of portfolio theory to regional decisions predicts the past. However, Spelman (2006) argued a portfolio perspective provides measures of current and expected future performance, enabling regions to choose to maximize growth or stability, based upon their own goals or preferences. Agents in the region itself are best acquainted with the region’s own employment trajectory, and are
thus also most likely to form their perceptions of future trends based on past region-specific experiences.

2.2. How Regions Affect Their Employment Portfolio

Since the 1970s employment portfolio analysis has been used to attempt to arrive at an acceptable balance between growth and stability (Jackson, 1984). For example, firm recruitment is influenced by the current portfolio of employers—diversify the employment portfolio or specialize and risk high volatility if the industry contracts. By focusing on growth and ignoring stability, regions increase fiscal, economic, and social risks (Barry & Kearney, 2006). Volatile industries are more likely to exhibit negative growth during a general economic downturn, while the most stable economies avoid this pain, but at the cost of decreased short-term employment growth. For this reason, “we can reasonably expect that citizens will want to consider both stability and growth in making economic development decisions,” (Spelman, 2006, p. 313).

This paper shows employment portfolios exist for smaller units of observation than previous research has found and that these employment portfolios affect entrepreneurship levels. That regions can affect their employment portfolio merely suggests that entrepreneurship might be used to affect a region’s employment portfolio, with the latter question remaining a topic for future research.

2.3. Units of Observation

Until now, employment portfolio studies were conducted at a large scale, i.e., states and metro areas, making local policy application of the work difficult (Lande, 1994). Application to state and metro areas is attractive because more data are available (Carlino, DeFina, & Sill, 2004; Conroy, 1974; Kort, 1981; Spelman, 2006). State-level
research, however, incorporates many diverse labor markets, potentially hampering regional application of results, e.g., Barth, et al. (1975) and Wundt (1992). Metro areas are more defensible labor market units than states, e.g., Spelman (2006), but omit surrounding rural areas as potentially idiosyncratic sources of both labor supply and labor demand. Thus, commuting sheds which incorporate rural areas may be the most appropriate areal unit (Jackson, 1984). Chandra (2003), however, found European provinces (NUTS 3) are too small a unit of analysis to successfully conduct regional portfolio analysis.

We find that U.S. counties and Commuting Zones (CZs) can be used for regional employment portfolio analysis.¹ CZs in particular represent an improvement over states and metropolitan areas as they are a more meaningful and plausible representation of local labor market agglomerations across both metropolitan and non-metropolitan/rural areas based on the commuting shed concept (Tolbert & Sizer, 1996). This paper is the first known successful application of Commuting Zones in the regional employment portfolio analysis literature, and the first successful sub-state portfolio analysis including all U.S. counties.

In addition to CZs, the availability of county-level data and commuting patterns based on the county-seat make counties another good choice for the unit of observation in this study. Counties are not just administrative units; many counties are separable labor markets, as sparsely populated rural economies are also often based on the county seat, generally the largest town in the region and usually centrally located geographically. Figure 1 shows CZs and counties respectively, illustrating the areal unit size difference.
2.4. Measuring Risk and Return

We measure growth, or return, as the average of annual wage and salary employment growth over the most recent business cycle and risk as the standard deviation of growth over the business cycle. We measure risk and return over a business cycle, trough-to-trough, because risk and returns change with regional structural shifts over multiple business cycles (Spelman, 2006). Several studies examine growth and volatility over multiple and/or incomplete, business cycles, but we use only the most recent complete business cycle because it is relevant for the entrepreneurship model, presented in section 3, and we want to focus on the most recent employment portfolio. Regardless, results of this analysis held for all business cycles for which data are available. We measure risk and return over the 1991-2001 business cycle.\(^2\) Data are obtained from the annual Bureau of Economic Analysis’ Regional Economic Information System.

Returns to a regional employment portfolio are measured by the overall growth of job prospects, which is similarly in the interest of both workers and firms. Most related literature uses employment growth as the preferred measure of return given its parallel with the underlying risk metric and the symmetric incentives for regional firms and workers (Barth, et al., 1975; Chandra, 2002, 2003). We measure return as the average of annual wage and salary employment growth.
Regional employment portfolio studies use labor market risk as a measure of volatility of returns (Lande, 1994). Local workers (and employers) appreciate the greater certainty of stable growth. Such stability can translate into a more attractive environment for developing high-skill labor pools at relatively lower wages, as workers are willing to trade some wages for reductions in employment risk. Firms also appreciate such stability for the clearer planning horizons it provides, as well as the labor pool it can provide (Krugman, 1991). In parallel to returns, we measure risk as the standard deviation of annual wage and salary employment growth over the 1991-2001 business cycle.

Like most employment portfolio literature, we use employment-based measures of risk and return rather than income-based because using employment avoids spurious results due to regional differences in cost-of-living. More directly, employment is beneficial to the firm, workers, and the region as a whole as it signals greater productive capacity. The divergence of perspectives between firms (that appreciate lower wages, ceteris paribus) and workers (who appreciate precisely the opposite dynamic) reinforces the utility of a focus on employment growth. Finally, trends in employment tend to broadly follow income trends and other measures of return to a region, such as population and tax revenue, leading to our use of employment as the focal measure of return without loss of generality (Barth et al., 1975). In any case, we find the general relationship between risk and return is substantively identical whether measured by income or employment.³

We use wage and salary employment, rather than total employment, because it is the most direct measure of opportunities in the formal labor market at a point in time, while also being unlikely to be a driver of the entrepreneurship measure introduced in
section 3 of this paper. The lack of such wage and salary employment opportunities—or volatility in such positions—increases the incentive to seek out self employment in the future, a proposition which is formally tested. Wage and salary employment excludes those who are self employed, allowing us to assess the incentive to engage in entrepreneurial self employment as a reaction to prospects in the formal wage and salary sector.

2.5. Model

The model assumes regions (and workers) are risk averse. A worker will not take a job if another job, or entrepreneurial opportunity, has more favorable levels of risk and return. Similarly, a region will prefer to recruit firms with the goal of reaching favorable levels of risk and return—or they may opt to funnel their recruitment dollars into fostering entrepreneurship or other economic development policies and programs.

Before we can test whether the risk-return tradeoff affects entrepreneurship, we must establish the existence of the risk-return tradeoff for counties and CZs. Prior research established that the return-risk frontier is U-shaped for states and metro areas and that the actual employment portfolios are associated with greater volatility is optimal—at a given time more regions are higher (more volatility) and more to the left (low growth) on the U-shape frontier than necessary (Lande, 1994; Spelman, 2006). Simply plotting growth and risk for U.S. counties and CZs over the 1991-2002 business cycle weakly suggests a U-shape relationship (Figure 2).
We utilize stochastic frontier estimation to test for the quadratic relationship between risk and return (Chandra, 2002; Jondrow, Knox, Lovell, Materov, & Schmidt, 1982). The stochastic frontier estimation method permits a test of the nonconvexity of the employment portfolio frontier and estimates the parameters that define its shape. The alternative, a nonparametric method called data envelopment analysis, will produce a U-shape frontier by construction (Spelman, 2006).

Stochastic frontier estimation is based on the assumption that the observations for each region lie on or within a frontier, creating one-sided disturbances. Aigner, Lovell, and Schmidt (1977) suggest two distributions for the one-sided disturbance, and we use the half normal distribution because its specification best suits regional portfolio applications of stochastic frontier analysis. We specify the dependent variable and regressors according to Chandra (2002); \( \sigma_i \) is the standard deviation of the rate of growth for region \( i \) over a business cycle, \( G_i \) is the rate of growth for region \( i \) over the business cycle, and \( \beta_1 \) and \( \beta_2 \) are the parameters to be estimated. We specify the error term, \( \epsilon_i \), according to Jondrow et al. (1982). The estimated equation is:

\[
(1) \quad \sigma_i = \alpha_i + \beta_1 G_i + \beta_2 G_i^2 + \epsilon_i, 
\]

where \( \epsilon_i = \nu_i - \mu_i, \nu_i \sim N(0, \sigma^2_{\nu}) \) and \( \nu_i \sim N(0, \sigma^2_{\nu}) \), hence, the half-normal distribution.

Except for the error term, OLS results are unbiased and consistent. A nonlinear estimator, the maximum likelihood estimator, however, is more efficient than OLS, thus, we use MLE to estimate equation (1), using functions developed in FRONTIER 4.1., by Tim Coelli (see, Battese & Coelli, 1988). Empirically, we estimate equation (1) using the variables Risk for \( \sigma_i \) and Growth for \( G_i \).
2.6. Results

The empirical test confirms that there is a U-shaped tradeoff between risk and return in county and CZ employment portfolios (Table 1). Plots confirm the global U-shape. These results affirm work by Spelman (2006), Chandra (2002), and Lande (1994) who also found a quadratic relationship between growth and volatility in income or employment. That this relationship exists suggests we can use risk and return to test whether employment portfolios affect entrepreneurship levels.

To the best of our knowledge, this analysis has only been successfully conducted at the sub-state level for metropolitan areas. Our finding is in opposition to Chandra (2003) who did not find statistically significant results for the stochastic test using European Union NUTS 3 (province) data.

3. Entrepreneurship Incentives

This section explores the likely supply-side influence of wage and salary employment portfolios, established in section 2, on engagement in entrepreneurial activity, measured empirically with self employment and establishment birth rates, by empirically testing whether employment portfolios affect next period’s entrepreneurship levels. Economic development professionals and policymakers often fail to understand the region’s labor market behavior when developing business recruitment and entrepreneurship development policies and programs. Better understanding how risk and
return incent nascent entrepreneurs will help with the implementation of better and more cost-effective programs to encourage economic growth and entrepreneurship.

3.1. Entrepreneurship and Local Labor Markets

We know of no research explicitly linking self employment or entrepreneurship and local employment portfolios. Loveridge and Nizalov (2007) found a region needs to maintain a sufficiently large share of self employed entrepreneurs and other small businesses to enhance growth in other parts of the economy, suggesting the importance of the self employed to the entire regional economy.

Parker (1996) provides motivation for our hypothesis. Using optimal control, he develops a model of self employment among heterogeneous agents who maximize discounted expected utility under uncertainty. The model predicts that paid employment opportunities affect the decision to become an entrepreneur and that risk and risk aversion negatively affect the probability an agent becomes self employed.

Parker finds evidence of people being pushed into self employment due to lack of wage and salary job opportunities, often during recessions. This push into self employment likely corresponds to the upper-left of the employment portfolio when high wage and salary employment volatility and low growth conditions exist (see, figure 2). Parker also finds evidence of people being pulled into self employment by high growth, developing markets, and increased disposable income; likely regions on the right of the U-shaped employment portfolio.
Both bodies of research hint at the explicit relationship between formal labor market risk and return, and the inherent incentives for both individual entrepreneurs and regional economies to value entrepreneurial business activity.

3.2. Defining Entrepreneurship for this Study

Defining entrepreneurship, "one of the most intriguing but equally elusive concepts in economics," (Peneder, 2009, p. 77), is a longstanding problem. A commonly accepted definition of entrepreneurship has failed to emerge (Gartner, 1988; Gartner & Shane, 1995; Luger & Koo, 2005) and defining and empirically measuring entrepreneurship, particularly at sub-state units of observation, is challenging and the topic of on-going research.

Self employment is a measurable at sub-state units and is the most widely used proxy for entrepreneurship (Blanchflower & Oswald, 1998; Glaeser, 2007), particularly for economic development and regional research (Iversen, Jorgensen, & Malchow-Moeller, 2008). Because a definition of and measure for entrepreneurship does not exist, we adopt self employment in-part because proprietors satisfy the basic characteristic of entrepreneurs: owner-management and risk bearing (Henderson, Low, & Weiler, 2007). Further, self employment is a suitable proxy for entrepreneurship for the purposes of this paper because it succinctly captures movement out of the formal wage and salary job market into self employment, the two being exclusive of each other. Despite its wide use, self employment is an imperfect measure of entrepreneurship. Self employment is overly broad, failing to distinguish between innovative and non-innovative activity, e.g., nascent high-tech firms and lifestyle businesses. County-level data on innovative
entrepreneurship does not yet exist, but is the topic of on-going research (see, Low, 2009).

As a robustness check, we also use establishment births as a metric for entrepreneurship. Establishment births is another widely used proxy for entrepreneurship (Acs & Mueller, 2008; Armington & Acs, 2002; Lee, Florida, & Acs, 2004) and some consider it a better measure of entrepreneurial activity than self employment (Acs & Armington, 2003). Timely, regional data on establishment births, however, are more difficult data to obtain and births do not capture individual's movement into and out of the formal wage and salary job market like self employment does, making coefficients of the establishment birth model more difficult to interpret.

Our principal dependent variable in the entrepreneurship model is the number of nonfarm proprietors normalized by total nonfarm employment, 2002. We also use the average nonfarm proprietorship rate 2002-2005 to ensure results are not specific to 2002. We measure the birth rate as new single-unit employer establishments, 2002 to 2003, normalized by 2002 employment.

3.3. Units of Observation

We model entrepreneurship at both the county and Commuting Zone level. County-level studies of entrepreneurship are becoming more common (Goetz & Rupasingha, 2009; Henderson, et al., 2007) but, to the best of our knowledge, no entrepreneurship studies have been conducted using Commuting Zones. Lee et al. (2004) find different factors drive new firm formation in Metropolitan Statistical Areas (MSAs) and Labor Market Areas (LMAs), likely due to the exclusion of the hinterland in LMAs, which suggests rural areas do influence the urban economy and should not be omitted.
from analysis, as is the case when MSAs are the unit of observation. We extend existing analysis of entrepreneurship by using commuting sheds that are less aggregated than those used by Lee et al. and Armington and Acs (2002) but still provide a consistent set of local labor market agglomerations across both urban and rural areas.

Using smaller units of observation will improve the entrepreneurship model, due to entrepreneurship’s pronounced regional and spatial dimension, which is partly due to entrepreneurial attitudes, financial capital, and human capital varying across space (Bull & Willard, 1993; Fritsch & Schmude, 2006). Additionally, most economic development is conducted at the local level (Bartik, 1991; Wasylenko, 1997), so, like employment portfolios, entrepreneurship policies should vary between regions.

3.4. Test and Model

We hypothesize the supply-side influence of wage and salary employment affects self employment, rates. We test this hypothesis by including employment portfolio components in an entrepreneurship model and testing if risk and return coefficients are different from zero. As Parker’s 1996 model suggests, we expect to find higher rates of self employment where wage and salary employment volatility is high—thus we expect a positive coefficient on risk. The coefficient on return, or employment growth, could vary because the relationship between entrepreneurship and growth is quadratic. Self employment is highest where growth is very low, e.g. entrepreneurship by necessity (Low, Henderson, & Weiler, 2005), but also where growth is very high, due to thick markets, more income, and the associated demand-side opportunities for entrepreneurs (Goetz & Rupasingha, 2009; Parker, 1996). We expect a positive coefficient on growth
squared, that is where growth is negative (far left on the U) or very high (far right on the U), entrepreneurship is higher.

3.4.1. Control Variables. We test the hypothesis (above) using an entrepreneurship model containing a vector of explanatory variables widely used in entrepreneurship models and an additional vector of employment portfolio variables. We borrow heavily from Goetz and Rupasingha’s (2008) vector of explanatory variable, X, so we can focus the analysis on the inclusion of the employment portfolio variables, R, as determinants of entrepreneurship. Like others, we include demographic and regional characteristics (Acs, Armington, & Zhang, 2007; Goetz & Rupasingha, 2009; Lee, et al., 2004).⁴ Demographic characteristics include financial capital and human capital. Regional characteristics include place-based characteristics and income-based characteristics. All variable descriptions, sources, and summary statistics are available in Appendix A.

Financial capital variables attempt to assess the ability of nascent proprietors to finance a new venture. The percentage of home owners and median home value in a region are proxies for the ability to self finance (Goetz & Freshwater, 2001; Goetz & Rupasingha, 2009). Bank deposits per capita, has been used by others to gain insight on the availability of financial capital, particularly in rural areas (Garofolli, 1994; Goetz & Rupasingha; Low, et al., 2005) and is a measure of how much money local banks have on hand for small business loans. Human capital control variables include percent of adults who are college educated and percent of non-college graduates who did not graduate from high school or receive a GED. These human capital variables have been used to
assess the relationship between human capital and entrepreneurship by Evans and Leighton (1989), Goetz and Freshwater, Lee et al. (2004), and Goetz and Rupasingha.

The percent of persons residing in rural areas, or outside places with fewer than 2,500 people, controls for the strong relationship between proprietorship and rural regions. The Natural Amenity Scale (McGranahan, 1999) is an index of topography, weather, and water, and is expected to control for the high levels of proprietorship in regions which are home to many “lone eagle” entrepreneurs who are attracted to the region by natural and scenic amenities (Carlino, Chatterjee, & Hunt, 2007; Kim, Marcouiller, & Deller, 2005).

We also include average wage and salary income and wage and salary income growth, rather than per capita income and per capita income growth, as Goetz and Rupisingha (2009) do. We use wage and salary income so that the focus remains on the wage and salary employment portfolio’s effect on the proprietorship rate. Wage and salary income captures the possibility that where wages are high entrepreneurs may be drawn to participation in the formal job market—the opportunity cost of self-employment is higher. Further, wage and salary income is a better measure of income because it is earned income and, unlike per capita income, excludes transfer payments, healthcare benefits, etc.

Regressors are represented by average county-wide or CZ-wide values so that each county or CZ proxies for the average characteristic of the potential pool from which proprietors are drawn (Goetz & Rupasingha, 2009).

3.4.2. Adding Employment Portfolio Variables to the Model. We add the employment portfolio variables to the entrepreneurship model to test the hypothesis that
conditions in the wage and salary job market affect the level of entrepreneurship. *Growth* is defined as wage and salary employment growth over the 1991-2002 business cycle and we also use its square. *Risk* is defined as the standard deviation of employment growth over the 1991-2002 business cycle. *Growth*, its square, and *Risk* compose the vector of employment portfolio explanatory variables which is denoted in equation (2) as the \( \mathbf{R} \) vector.

3.4.3. Estimation Issues. We use lagged explanatory variables to reduce the endogeneity bias, show that at least some of the causality flows from risk/return to proprietorship, and allows current entrepreneurship to be the longer-term result of local workers considering and acting upon a region’s formal wage and salary employment prospects over time. The estimated equation is:

\[
Y_t = \beta \mathbf{R}_{t-1} + \delta \mathbf{X}_{t-1} + \varepsilon_{t-1}.
\]

There is likely still endogeneity in the model due to the error term containing drivers of entrepreneurship and education; this simultaneity will cause results to be biased, but consistent. Nevertheless, we estimate the linear model for counties and CZs using 2SLS (not shown) and found the results to be robust in coefficient sign and significance. Multicollinearity is not problematic in the model; the variance inflation factor, VIF, for each explanatory variable is below five. Equation (2) estimated by OLS suffers from heteroskedasticity of the error terms.\(^5\) The White-Huber correction is applied and standard errors re-estimated; results (not shown) remain consistent despite the correction.

3.4.4. Spatial Issues. Heteroskedasticity in OLS models is one of the first indicators that the errors contain a spatial process. We suspect spatial processes likely
exist because businesses, markets, and the labor force vary widely by region (Klein & Cook, 2006). Chinitz (1961) argues that entrepreneurship supply is a function of society, capital availability, and willingness to bear risk—all of which vary regionally. By controlling for spatial processes, noise is reduced, model fit is improved, and, importantly, heterogeneity is controlled for (Anselin, 1988).

Spatial error processes, or nuisance errors, can occur due to omitted spatially correlated variables or the value of adjacent observations moving together due to common or correlated unobservable variables. Experience tells us the county model likely suffers from spatial error but developing a theory behind implementation of the spatial error model is difficult because the errors are not due to some underlying process—rather a host of micro processes. To improve model efficiency beyond OLS with a White-Huber correction, we use the spatial error model (SEM) to control for spatially correlated disturbances between counties:

\[ Y = \beta X + \varepsilon, \]

where, \( \varepsilon = \lambda W \varepsilon + u, \) \( u \) i.i.d.

Spatial lag processes occur due to interaction among neighbors. The use of the spatial lag model for Commuting Zones is appealing because CZs are small enough that networking and interaction between neighbors exists, but big enough that nuisance errors do not override the spatial lag process, as is common in county-level models. When spatial lag processes are unaccounted for, coefficients can be biased and inconsistent—leading to the wrong sign on coefficients or invalid hypothesis testing. Spatial dependence of this type is most frequently incorporated into models using the spatial autoregressive (SAR) lag model:
\( Y = \rho W Y + \beta X + \varepsilon \),

where \( \varepsilon \sim \text{i.i.d.} \), and \( W \) is an \( n \times n \) matrix defining spatial unit interaction (Anselin, 1988).

There is very little formal guidance in the choice of the optimal spatial weights matrix (Anselin, 2008b). We use a first order queen contiguity weights matrix due to the nature of spatial dependence and its suitability for use with irregular polygons. The weights matrix is row-standardized to facilitate with interpretation and ease computational expense.

Lagrange Multiplier (LM) tests are frequently used to test whether the spatial error or spatial lag process is present. For U.S. counties we find the lag and error LM tests are significant, thus we conduct robust LM tests. The robust LM test (error) is significantly different from zero and the robust LM (lag) test is insignificant, thus, the robust LM tests show preference for the SEM, confirming our hypothesis that nuisance errors exist in the U.S. counties model. For CZs both LM tests are again significant, so we run the robust LM tests which suggested the spatial lag model is dominant over the spatial error for CZs, likely because there are less nuisance errors due to fewer and larger spatial units. Results of the LM tests affirm the expectations discussed; the CZ model is best suited to a SAR model.

3.4.5. Estimated Model. The county model is estimated with the SEM model, equation (3), using an optimal generalized method of moments covariance matrix robust to unspecified forms of heteroskedasticity detected in the OLS model with the Breusch-Pagan test. The CZ model is estimated with a spatial autoregressive lag model, equation (4), using a maximum likelihood estimation.
3.5. Results and Policy Implications

3.5.1. Hypothesis Test Results. We find evidence to reject the null hypothesis; that is, wage and salary employment risk and return do affect a region’s self employment and establishment birth rates (Table 2). Significant coefficients on the variables of interest, risk, growth and its square, suggest that researchers should incorporate indicators of the regional employment portfolio into studies on the determinants of entrepreneurship. Finally, that employment opportunities and volatility affect the entrepreneurship level in a region suggests that wage and salary job opportunities (and industrial recruiting) must be considered in tandem with entrepreneurship development policies.

It is worth underscoring a cautionary note before proceeding with a discussion of the results. Due to estimation methods employed, estimated coefficients cannot be interpreted in the same straightforward manner as OLS results. For example, in a linear OLS estimation we can interpret a coefficient as the difference in the predicted value of Y, self employment, for each unit difference in X, say a standard deviation difference in Risk, ceteris paribus. This interpretation is more complex in spatial econometric models due to nonlinear estimation methods, the spatial multiplier effect and spatial connectivity—a change in an explanatory variable in region i will have a direct impact on region i as well as an indirect impact on other regions. For a detailed discussion of these complexities, we refer the reader to Aneslin (2003) or Le Sage and Pace (2009). Due to complications associated with interpreting coefficients in spatial econometric models, we focus rather on assessing whether results are reasonable and providing a general sense of the magnitude of test results.
Results of the test on \textit{Risk} align with our intuition that when wage and salary employment volatility is higher, more people are self employed in the next period—likely a response to labor market conditions. The effect of wage and salary employment volatility on self employment is relatively large. In the county-level model, a one standard deviation increase in risk increases the self employment rate by 63 percent (Table 2), an expectedly large result. Given the complications associated with interpreting spatial econometric coefficients, this magnitude must be considered judiciously; however, the linear OLS estimate was in a similar range at 70 percent. For commuting zones, the increase is 106 percent, likely because the spatial lag model, indicated for CZs, includes both direct and indirect effects. In short, test results suggest volatility in the wage and salary job market is associated with more workers being self employed in the next period—or being their own boss, whether out of necessity or opportunity.

Results suggest an inverse relationship between wage and salary employment growth and the self employment rate dominates (Table 2). The spatial econometric coefficient suggests the hypothesized relationships exist, but are relatively smaller than the \textit{Risk} results. In the CZ model, predicted self employment is 1.1 percent higher when wage and salary employment growth decreases by 1 percent, that is, when jobs decrease by 1 percent the next period's self employment rate is 1.1 percent higher. The square of wage and salary growth is positive and significant, as expected, suggesting the growth affect occurs at an increasing rate. Wage and salary employment growth was insignificant.
in the county-level model, likely because counties are a less appropriate areal unit than CZs. Growth squared is significant and positive, as expected, suggesting that the wage and salary employment rate decreases at an increasing rate.

Coefficient signs and significance are the same in the model of establishment births (Appendix B). However, coefficient magnitudes suggest that births are less responsive to wage and salary employment risk and growth in the previous period than is self employment. Interpretation of coefficients in the employer establishment birth model is less meaningful than in the self employment model because births do not capture individuals' movement between self employment and wage and salary employment as succinctly as self employment. Finally, tests using self employment for 2002-2005 yielded similar results, suggesting results are not specific to 2002.

3.5.2. Demographic Variables. The coefficient on the high school dropout rate was negative and significant (0.01), suggesting that efforts to decrease the high school dropout rate may increase proprietorship. Percent with a college degree was statistically insignificant, although others have found a positive and significant coefficient (Henderson et al., 2007). The test of homeownership as a proxy for capital availability was inconclusive, however, bank deposits per capita has a positive and significant (0.01) driver of entrepreneurship in both counties and CZs. This result suggests that efforts to improve access to small amounts of capital may foster entrepreneurship. An example of such would be establishing a revolving loan fund for nascent entrepreneurs in the region.

3.5.3. Regional Variables. Regional variables are place-based and income-based. The only place-based variable which has policy implications is the natural amenity scale, because rurality is difficult to influence with policy. In both regressions the coefficient on
natural amenities has a positive and significant (0.01) coefficient, which was expected because recent literature has found natural amenities and landscape to influence entrepreneurship (Henderson et al. 2007). Prior research has also suggested that amenities are necessary to attract and retain skilled professionals (Gale, 1998). Although amenities cannot be changed in the short term, regions can influence the appearance and accessibility of their environment. Mountaintop strip-mining may negate the amenity value of mountains in West Virginia. Planting trees, providing recreational access to buffer strips, and rails to trails projects may increase the amenity value of a place.

Income-based variables were insignificant in the CZ model, but we find income and income growth to be negatively associated with the proprietorship rate in the county model at the 0.01 level of significance. This could be because the nonfarm proprietor income is below wage and salary income.

4. Conclusions

We find that risk and return in the wage and salary job market affect self employment and, to a lesser extent, establishment birth rates, in the next period. In uniting two previously disparate literatures, this paper makes contributions to both the entrepreneurship and employment portfolio literature as they relate to local labor market research and development. The paper is the first known to connect regional employment portfolio theory with entrepreneurship, particularly in terms of the incentive to pursue self employment given the risk/return patterns of local wage and salary positions. Second, this paper represents the first consistently successful application of portfolio theory to sub-metropolitan economies such as counties and CZs, with commuting zones in
particular allowing both broader and deeper insights into local labor market interactions through their commuting shed agglomerative lens. Furthermore, by better understanding the determinants of entrepreneurship and realizing that wage and salary employment opportunities can affect entrepreneurship incentives, policymakers can better foster entrepreneurship in their region. Entrepreneurial clubs (Green, Wise, & Armstrong, 2007), workshops and training (Camp, 2005), financial capital (Garofolli, 1994), and entrepreneurial development systems (Lyons & Hamlin, 2001) may increase entrepreneurship incentives. A better understanding of how job prospects affect entrepreneurship may also provide information to better allocate economic development dollars between recruitment and entrepreneurship.

The spatial detail of this research also empowers regional decision making—rather than all regions following a blanket policy (Klein & Cook, 2006). If a region is above and left of the optimal frontier, leaders can allocate scarce economic development dollars differently with the goal of reducing volatility and/or increasing employment growth. This result is important because a lot of economic development policy in the United States is conducted at the local level—meaning, for the first time, portfolio analysis of labor markets can be used by economic development practitioners rather than having to use less-specific state results.

This research begets more research. Focusing on the differences in entrepreneurship incentives between urban and rural counties and CZs may explain some of the results of this paper. Future research will examine whether entrepreneurship can be used to decrease volatility and/or increase growth while incorporating industry sectors into the analysis—the idea being that regions encourage entrepreneurship in sectors
which behave counter-cyclically to existing industry, which would further decrease regional employment portfolio volatility.
Notes

1 Tolbert and Sizer (1996) define two areal units which encompass regional labor markets: Commuting Zones (CZ), n=741, and more aggregated Labor Market Areas (LMA), n=394, both of which can cross state lines, are the product of a consistent methodology, and include rural counties (Hammond & Thompson, 2004). Tolbert and Sizer aggregate counties by commuting patterns and find that CZs are meaningful spatial units and plausible representation of local economies. CZs are particularly useful because there is no minimum population threshold, unlike the LMAs, which were designed to meet the Census Bureau’s criterion of a 100,000 population minimum. For these reasons, we use CZs rather than LMAs.

2 The NBER (http://www.nber.org/cycles/cyclesmain.html) trough-to-trough business cycle is March 1991 to November 2001. When we plotted the annual wage and salary employment data, the trough years were 1991 and 2002, thus we use 1991-2002 in the empirical exercise.

3 The U-shape tradeoff between average income and income volatility is statistically significant, suggesting both income and employment growth are robust proxies for return. Additionally, income-calculated risk and return were both plotted and found to have a U-shaped relationship, globally.

4 We do not include Goetz and Rupasingha’s (2009) policy characteristics because data are only available at the state-level.

5 The Breusch-Pagan (BP) test for heteroskedasticity rejects the null hypothesis of no heteroskedasticity in the error terms BP=285.6, p <0.0001
Often, both spatial error and lag processes are present, but detecting the presence of both is difficult because the test tends to be significant when either the error or lag alternative hypothesis is proper, but not necessarily both (Anselin, 2008b). The LM test with SARMA(1,1) alternative hypothesis suggests a higher order alternative model with both a spatial error and a spatial lag term is appropriate, but rejection of the null does not necessarily imply that the higher order model is the proper alternative. Additionally, the use of both spatial error and spatial lag terms in econometric models often leads to identification problems and it is likely that a different specification of the weights will eliminate the need for a higher order model which incorporates both spatial lag and spatial error (Anselin, 2008a).
References


### Appendix A: Variable Description and Summary Statistics

<table>
<thead>
<tr>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>0.032</td>
<td>0.027</td>
<td>0.006</td>
<td>1.06 Standard deviation of annual wage and salary emp. growth, 1991-2002</td>
</tr>
<tr>
<td>Growth</td>
<td>0.014</td>
<td>0.018</td>
<td>-0.052</td>
<td>0.422 Average of annual wage and salary employment growth, 1991-2002</td>
</tr>
<tr>
<td>Growth squared</td>
<td>0.001</td>
<td>0.004</td>
<td>0</td>
<td>0.178 Growth (see above), squared</td>
</tr>
<tr>
<td>% college educated</td>
<td>0.133</td>
<td>0.063</td>
<td>0.037</td>
<td>0.534 Percent of adults &gt;25 years with a 4-year college degree</td>
</tr>
<tr>
<td>% high school dropouts</td>
<td>0.346</td>
<td>0.103</td>
<td>0.074</td>
<td>0.734 Percent of adults &gt;25 years without a 4-year degree who have not</td>
</tr>
<tr>
<td>Bank deposits per capita</td>
<td>36.6</td>
<td>213.2</td>
<td>0</td>
<td>6875.8 Bank deposits ($1000) per capita, FDIC, 1990</td>
</tr>
<tr>
<td>Average income</td>
<td>17.34</td>
<td>3.713</td>
<td>9.829</td>
<td>40.526 Wage and salary income ($1000) divided by population, 1990</td>
</tr>
<tr>
<td>Natural amenities</td>
<td>0.055</td>
<td>2.313</td>
<td>-6.4</td>
<td>11.17 McGranahan's (1999) amenity scale, USDA ERS</td>
</tr>
<tr>
<td>% of home owners</td>
<td>0.618</td>
<td>0.091</td>
<td>0.101</td>
<td>0.827 Percent of residences which are owner occupied</td>
</tr>
<tr>
<td>Median home value</td>
<td>52,353</td>
<td>32,107</td>
<td>14,999</td>
<td>500,001 Median home value ($)</td>
</tr>
<tr>
<td>% residing in rural place</td>
<td>0.637</td>
<td>0.291</td>
<td>0</td>
<td>1 Farm plus nonfarm rural residents divided by population</td>
</tr>
<tr>
<td>Income growth</td>
<td>0.753</td>
<td>0.564</td>
<td>-0.303</td>
<td>16.052 Wage and salary income growth, 1991-2002</td>
</tr>
<tr>
<td>Self employment rate</td>
<td>0.24</td>
<td>0.091</td>
<td>0.029</td>
<td>0.678 Nonfarm proprietor employment over nonfarm employment, 2002</td>
</tr>
</tbody>
</table>

*a* Bureau of Economic Analysis (Regional Economic Information System), 1991-2002, unless year is otherwise noted

*b* U.S. Decennial Census, STF3, (1990)
Appendix B: Results with alternative dependent variable: Establishment birth rate

The hypothesis test coefficients are the same in sign and significance as results using nonfarm self employment as the dependent variable. Summary statistics and a description of the dependent variable are in Appendix A.

Table Appendix B: Regression Results, Y=births, for County and CZ

<table>
<thead>
<tr>
<th></th>
<th>County Coefficient (error)</th>
<th>Std Error</th>
<th>CZ Coefficient (error)</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>15.664*** (1.234)</td>
<td>0.008***  (0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk</td>
<td>47.741*** (6.974)</td>
<td>0.052**   (0.037)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>-2.770 (9.906)</td>
<td>-0.032*   (0.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth squared</td>
<td>1169.0*** (151)</td>
<td>1.314***  (0.418)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent college educated</td>
<td>7.573*** (2.484)</td>
<td>0.004     (0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct high school dropouts</td>
<td>-8.117*** (1.204)</td>
<td>-0.003**  (0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank deposits per capita</td>
<td>0.0008** (0.0005)</td>
<td>0.00007*** (0.00002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of home owners</td>
<td>-3.341*** (1.202)</td>
<td>-0.010*** (0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median home value</td>
<td>0.00002 (0.000004)</td>
<td>0.00000001 (0.0000007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural amenities</td>
<td>0.640*** (0.050)</td>
<td>0.0005*** (0.0005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pct residing in rural place</td>
<td>6.656*** (0.440)</td>
<td>0.006*** (0.0005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average income</td>
<td>-0.382*** (0.036)</td>
<td>-0.00000005 (0.000005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income growth</td>
<td>-4.184*** (0.511)</td>
<td>-0.001    (0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ (LR test value 23.3)</td>
<td>0.148***</td>
<td>0.331***  (0.040)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ</td>
<td></td>
<td>0.331***  (0.040)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log L</td>
<td>23.3</td>
<td>3342.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.I.C.</td>
<td>17985.0</td>
<td>-6654.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nc</td>
<td>2959</td>
<td>720</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significant at *10%, **5%, ***1%

* N is lower than total number of counties and CZs due to missing data.
Table 1: Test for U-Shape Risk/Return Tradeoff: Counties and CZs\textsuperscript{d}

<table>
<thead>
<tr>
<th></th>
<th>CZ Estimation</th>
<th></th>
<th>County Estimation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std Error</td>
<td>Estimate</td>
<td>Std Error</td>
</tr>
<tr>
<td>Constant</td>
<td>0.012***</td>
<td>(0.0005)</td>
<td>0.011***</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.410***</td>
<td>(0.041)</td>
<td>-0.230***</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Growth squared</td>
<td>12.9***</td>
<td>(1.437)</td>
<td>6.6***</td>
<td>(0.115)</td>
</tr>
<tr>
<td>N\textsuperscript{c}</td>
<td>720</td>
<td></td>
<td>N\textsuperscript{c}: 2959</td>
<td></td>
</tr>
</tbody>
</table>

LR test of the one-sided error = 174.41
log likelihood value: 2499.401
Significant at ***1%

LR test of the one-sided error = 899.14
log likelihood value: 8271.677

\textsuperscript{c} N is lower than total number of CZs and counties due to missing data
\textsuperscript{d} For variable description and summary statistics see Appendix A
<table>
<thead>
<tr>
<th></th>
<th>County Coefficient</th>
<th>CZ Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td>0.324*** (0.018)</td>
<td>0.151*** (0.026)</td>
</tr>
<tr>
<td><strong>Risk</strong></td>
<td>0.631*** (0.101)</td>
<td>1.062** (0.616)</td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td>-0.188 (0.144)</td>
<td>-1.142*** (0.315)</td>
</tr>
<tr>
<td><strong>Growth squared</strong></td>
<td>7.288*** (2.199)</td>
<td>21.0*** (6.990)</td>
</tr>
<tr>
<td><strong>Percent college educated</strong></td>
<td>-0.026 (0.036)</td>
<td>-0.022 (0.049)</td>
</tr>
<tr>
<td><strong>Pct high school dropouts</strong></td>
<td>-0.197*** (0.018)</td>
<td>-0.095*** (0.020)</td>
</tr>
<tr>
<td><strong>Bank deposits per capita</strong></td>
<td>0.00001*** (0.000007)</td>
<td>0.002*** (0.0003)</td>
</tr>
<tr>
<td><strong>Percent of home owners</strong></td>
<td>0.061*** (0.017)</td>
<td>-0.081*** (0.026)</td>
</tr>
<tr>
<td><strong>Median home value</strong></td>
<td>0.0000005*** (0.00000007)</td>
<td>-0.0000001 (0.00001)</td>
</tr>
<tr>
<td><strong>Natural amenities</strong></td>
<td>0.005*** (0.001)</td>
<td>0.006*** (0.0008)</td>
</tr>
<tr>
<td><strong>Pct residing in rural place</strong></td>
<td>0.127*** (0.006)</td>
<td>0.123*** (0.008)</td>
</tr>
<tr>
<td><strong>Average income</strong></td>
<td>-0.009*** (0.001)</td>
<td>0.00004 (8.882)</td>
</tr>
<tr>
<td><strong>Income growth</strong></td>
<td>-0.045*** (0.007)</td>
<td>0.010 (0.021)</td>
</tr>
<tr>
<td>λ (LR test value 31.204)</td>
<td>0.160***</td>
<td></td>
</tr>
<tr>
<td><strong>ρ</strong></td>
<td></td>
<td>0.257*** (0.042)</td>
</tr>
<tr>
<td>Log L</td>
<td>3546.4</td>
<td>1317.0</td>
</tr>
<tr>
<td>A.I.C.</td>
<td>-7060.8</td>
<td>-2602.0</td>
</tr>
<tr>
<td>N c</td>
<td>2959</td>
<td>720</td>
</tr>
</tbody>
</table>

Significant at **5%, ***1%

c Observations with missing data are dropped
Figure 1a: Commuting Zones

Figure 1b: U.S. Counties
Figure 2: CZ Risk and Return, 1991-2001 Business Cycle.