Entrepreneurship, Information, and Growth

Devin Bunten[1,*], **Stephan Weiler**[2,†], **Stephan Weiler**[3,‡], and **Sammy Zahran**[4,§]

[1]University of California, Los Angeles

[2]Colorado State University

[3]University of Nebraska-Lincoln

[4]Colorado State University

Abstract

We examine the contribution to economic growth of entrepreneurial "marketplace information" within a regional endogenous growth framework. Entrepreneurs are posited to provide an input to economic growth through the information revealed by their successes and failures. We empirically identify this information source with the regional variation in establishment births and deaths, which create geographic information asymmetries that influence subsequent entrepreneurial activity and economic growth. We find that local establishment birth and death rates are significantly and positively correlated with subsequent entrepreneurship for US counties. To account for the potential endogeneity caused by forward-looking entrepreneurs, we utilize instruments based on historic mining activity. We find that the information spillover component of local establishment birth and death rates have significant positive effects on subsequent entrepreneurship and employment growth for US counties and metropolitan areas. With the help of these intruments, we show that establishment births have a positive and significant effect on future employment growth within all counties, and that in line with the information hypothesis, local establishment death rates have a similar positive effect within metropolitan counties.

1 Introduction

In this paper, we explicitly quantify and incorporate entrepreneurship into a spatial-equilibrium endogenous growth framework to better understand the role of entrepreneurs in determining economic growth. We emphasize the revelation of marketplace information and a resultant externality, aspects of entrepreneurship that have not been identified previously in the literature. Through the successes and failures of their projects, entrepreneurs generate valuable marketplace information regarding the contours of the geographic and industrial territory in which their projects reside. This externally beneficial information can be utilized by future entrepreneurs, who can emulate successful projects and avoid the pitfalls identified by the failures.

[*]devinbunten@ucla.edu.
[†]stephan.weiler@colostate.edu.
[‡]ethompson2@unl.edu.
[§]szahran@colostate.edu.

We quantify this understanding by focusing not on the entrepreneur but on her project. To undertake a project, an entrepreneur must assess local and broader demand for her products and services, the necessary supply network, and the feasible financing. Each entrepreneurial action—opening, expanding, or even terminating a project—illuminates a niche of the marketplace. Project financiers must assess these same aspects as well as the suitability of the entrepreneur herself. Through its evolution and possible eventual demise, the project itself provides information on the viability of similar projects.

Our basic hypothesis is maintained empirically. Spatial differences in entrepreneurial activity, as expressed through the births of local establishments, have a statistically and economically significant effect on local employment growth. This finding is robust to a variety of specifications. Furthermore, local establishment births *and deaths* are strongly associated with future establishment birth rates—evidence that future entrepreneurs look to past successes *and failures* when choosing what projects to implement. Finally, we close the loop by identifying twin indirect positive impacts that establishment deaths have on future employment growth within metropolitan counties—effectively underscoring the likely informational role of entrepreneurial projects on growth in these populous counties. We find similar results when we include establishment births alongside existing measures of entrepreneurship; these results support our core contention that establishment births capture a relationship with employment growth that is different from its relationship to alternatives such as proprietorship and small business employment.

The fundamental relationship between entrepreneurial projects and employment growth remains consistent when using instrumental variable analysis featuring historical mining activity as a first-stage instrument. As in Glaeser et al. (2012), we follow the Chinitz (1961) hypothesis that large-scale mining precipitated a large-scale industrial structure inimical to entrepreneurship. These results are robust to a variety of specifications of independent variables and remain remarkably strong even after including traditional controls known to influence local growth.

We proceed with a discussion of the existing literature and the motivation for studying entrepreneurial projects within an endogenous growth framework, and present the model. Then, we test the intermediate hypothesis that levels of entrepreneurship are positively dependent on past establishment births and deaths before proceeding to the main growth analysis. The results are highly supportive of our main hypothesis that entrepreneurship causally supports economic growth, and further evidence suggests that the marketplace information framework is useful for interpreting these findings. The paper closes with a discussion of these results and a brief conclusion.

2 Motivation and Background

Recent studies like Audretsch and Keilbach (2004, 2007) have provided substantive evidence of a relationship between entrepreneurship and growth, while Àcs et al. (2009) and Braunerhjelm et al. (2010) provide models featuring entrepreneurs who implement innovations flowing from spilled-over knowledge and research. These approaches update earlier endogenous growth models including those of Krugman (1987, 1991), Lucas (1988),

and Romer (1986, 1990) wherein knowledge or human capital are created endogenously and external effects of differential knowledge cause divergent outcomes across economies. Àcs and Armington (2006) emphasize entrepreneurial processes at the local level, echoing the hypotheses of Lucas (1988) and Jacobs (1970). Our paper corroborates these findings and provides evidence that revealed marketplace information is an important mechanism in accounting for the relationship between entrepreneurship and growth.

Information has a distinctly local characteristic in such a formulation (e.g. Weiler, 2006). The potential for business success is market-specific, both in its potential demand contours (e.g. for non-traded service industries) as well as more universally in its supply/cost character (e.g. labor, land, capital, insurance, among others). In regions with "thinner" markets—featuring fewer openings, closings, and other business transactions—information on the potential of such markets will consequently be more limited. While such information gaps may themselves impede prospective entrepreneurs directly, they may also indirectly restrict business opportunity through higher perceived uncertainty by critical loan and insurance suppliers. In the spirit of Akerlof (1970), we thus see differences in market-specific entrepreneurial experience as driving *geographic information asymmetries*, which can in turn yield suboptimal local investment and growth.

Previous work incorporates revealed marketplace information within a geographic context. Lang and Nakamura (1993) apply the geographic asymmetry logic to neighborhood housing markets where numerous transactions produce precise information, reduce future-period uncertainty, and encourage future activity. Weiler (2000a) provides a game-theoretic model and case study of information revelation where, upon the success of a pioneering firm, project viability is revealed and other firms make entry decisions. Weiler et al. (2006) extend the theoretical structure of entrepreneurial information revelation in a Bayesian Revealed Preference framework, where the perceived probability distribution of project outcomes is updated through discrete information increments.

The logic of geographic information asymmetries is borne out in the financial literature. Distance is a significant factor in determining abnormally high returns in firm acquisitions, with more localized transactions allowing greater insight into the most promising targets (Basu and Chevrier, 2011). These findings are particularly strong for relatively small, nonpublic, R&D-intensive, non-metropolitan firms with no analyst coverage—precisely the most informationally opaque to non-locals (Uysal et al., 2008). These high returns notwithstanding, Weiler et al. (2006) find that the social benefits of information provision are nevertheless likely to exceed the private benefits, suggesting a potential market failure.

Local context presents an opportunity for the study of economic growth amongst regions with broadly similar institutions and populations. International macroeconomists view integrated local economies as potentially revealing contexts for better understanding the microeconomic mechanisms of the macroeconomic growth process (e.g. Krugman, 1991), while labor economists are leveraging local labor market analysis in pursuit of this same lens (e.g. Moretti, 2011). Human capital and its localized external effects have been emphasized by Shapiro (2006), Moretti (2004), Glaeser et al. (1995) and Hammond and Thompson (2008). Such work is further motivated by the tremendous variation in economic

performance of subnational economies, a question which has challenged the economics field for decades (e.g. Hall et al., 1970; Summers et al., 1986; Blanchard et al., 1992; Weiler, 2000b). Low and Weiler (2012) in fact suggest that entrepreneurs take these variable conditions into account when choosing to become self-employed, as their decisions are directly influenced by the risk and returns of wage and salary employment opportunities in local labor markets.

Entrepreneurship may be one of the fundamental microeconomic mechanisms leading to geographically asymmetric outcomes—indeed, similar reduced-form analysis by Stephens and Partridge (2011) and Stephens et al. (2013) suggests that the lagging Appalachian region could best be boosted by a focus on entrepreneurship. We hold that the finding is general in that it applies nationally, and provide evidence that entrepreneurial activity and marketplace information specifically are critical to economic growth. The next section presents a model of spatial equilibrium and endogenous growth to aid in assessing whether entrepreneurial decisions to undertake, expand, and terminate micro-economic projects determine the path of local economic development.

3 Model and Data

We outline below a model of endogenous growth that is premised on the Roback (1982) model of a spatial equilibrium resulting from the maximizing choices of households and firms. Following Stephens et al. (2013), we consider labor migration flows and their determinants. For households, the key choices revolve around relative real wages and local amenities, while relative wages and productivity are determinative for firms. Structural factors combine with agents' decisions to produce equilibrium wages, house prices, and ultimately net labor flows; these flows become the object of our reduced-form analysis.

Households maximize utility U by choosing from a range of cities, indexed by i, the location that maximizes U based on the expected wages for the household's relevant skills, the prices of non-traded land and services, place-based amenities, and other local and idiosyncratic characteristics. Household choices of optimal locations will produce a long-run equilibrium in which the utility flows from locating in different regions are equalized for identical households. In the medium run, changes in local labor supply will depend positively on relative local utility ($U_i - U$, where U is their utility elsewhere) net of moving costs (M_i) as formalized in (1), where $L^S(.)$ is increasing in its argument.

$$\Delta Labor\ Supply_i = L^S(U_i - U - M_i) \quad (1)$$

Firms maximize profits π by choosing from a range of cities the location to produce that maximizes π based on an array of considerations: the local wage, the size and skill composition of the labor market, the availability and cost of local non-traded inputs (e.g. land), the area's access to markets both internal and distant, and other local productivity determinants including agglomeration economies. When choosing a location, firms weigh these factors and locate where expected profit is greatest. In the long run, entry and

relocation will equilibrate expected profits; in the medium run changes in local labor demand will depend positively on relative profits ($\pi_i - \pi$, where π is their profit elsewhere) as summarized in (2), where $L^D(.)$ is increasing in its argument.

$$\Delta Labor\ Demand_i = L^D(\pi_i - \pi) \quad (2)$$

Equations 1 and 2 can be combined to create a reduced form relationship between growth in the local labor market and the underlying determinants of household-side amenities and firm-side productivity effects. In the reduced form, local employment growth is a function of local economic characteristics that influence the growth of labor supply or demand. Characteristics that influence utility will affect growth in labor supply. Natural amenities influence utility, as does proximity to a metropolitan area given the variety of job openings and social and cultural opportunities afforded by cities. Likewise, local characteristics that influence profitability will affect growth in labor demand. Demand shocks related to the industries present in a region influence profitability as do characteristics that determine productivity such as the human capital of the local workforce and agglomeration economies resulting from the scale and scope of local economic activity.

Equation (3) shows the reduced-form relationship between local employment growth and its various determinants. Following the above discussion, these determinants include household-side amenities (A_i), measures of human capital including worker education (BA_i) and occupation profile (OC_i), idiosyncratic productivity effects or demand shocks (Z_i)—and others including the share of individuals in creative occupations (CC_i) and market access (MA_i).[1] To the list of employment growth determinants we append the propensity to engage in entrepreneurial projects (EP_i).

$$Growth_i = G(A_i, BA_i, OC_i, Z_i, CC_i, MA_i, EP_i) \quad (3)$$

Entrepreneurial projects are included in Equation (3) in accord with our hypothesis that entrepreneurship will increase productivity of the local economy through information spillovers. Any entrepreneurial project will likely involve management and resource allocation, innovation, financial risk-taking, and other tasks and chores.[2] Rather than emphasize any particular aspect of the entrepreneurial process, the experimentation involved in a project is the core of our theoretical understanding of entrepreneurship—whether that experimentation involves a new product or process, a new allocation of resources within an industry or region, or any other change from the previous economic arrangements of a locale. Such a project has the potential to create social benefits beyond its private benefits in the form of increased marketplace information. Through their success and failures,

---

[1]Physical capital—along with wages and other prices—is excluded here as maximizing agents will equilibrate its return across locales and so it cannot play a role in the reduced-form empirical implementation.
[2]Left untouched is the source of inspiration for entrepreneurial projects. Knowledge spillovers from research and development are one possibility, but we remain agnostic on the ultimate source of project inspiration.

entrepreneurial projects reveal the shape of market demand, supply conditions, and the viability of concepts or innovations from elsewhere to work within a particular context. They enable established firms, future entrepreneurs, and financiers to make better-informed decisions, whether they choose to replicate or support successful projects or to improve upon or avoid the pitfalls highlighted by unsuccessful ones. This improved information directly affects the productivity of enterprises within a locale.

Based on this project-oriented informational perspective, we identify establishment births and deaths as our key measure of entrepreneurship. An establishment birth aligns nicely with the theoretical understanding of an entrepreneurial project: both brand-new firms and new locations for expanding firms involve the experimentation highlighted above, and both induce the informational externalities at the core of this understanding of entrepreneurship. Establishment deaths generate information about types of entrepreneurial projects a local economy could not support. These metrics present an improvement on traditional measures such as the share of small firms or of the self-employed, both of which include many stagnant enterprises while neglecting the information content provided by establishment births and deaths.

As suggested by the reduced form growth model, we utilize the empirical strategy shown in Equation (4) below:

$$Growth_i = \beta_0 + \beta_1 * Entrepreneurial\ Projects_i + \beta_2 * Other\ Entrepreneurial$$
$$+ \beta_3 * Amenities_i + \beta_4 * Share\ with\ BA_i + \beta_5 * Share\ High\ Human\ Capital_i$$
$$+ \beta_6 * Demand\ Shock_i + \beta_7 * Age_i + \beta_8 * Density_i + \beta_9 * Income_i + \beta_10 * Employment$$
$$+ \beta_{11} * Distance\ to\ Metro_i + \beta_{12} * Lagged\ Employment\ Growth$$
$$+ \beta_{13} * Lagged\ Population\ Growth + \varepsilon_i$$
$$(4)$$

The control variables follow our reduced form model and existing literature including Stephens and Partridge (2011); Stephens et al. (2013) and also Hammond and Thompson (2008), Rappaport (2004, 2007), Partridge et al. (2009), Beeson et al. (2001), Àcs and Armington (2006), Huang et al. (2002), Deller et al. (2001) and Glaeser et al. (1995). The variables include education attainment, the share of employment in high human capital and creative occupations, natural amenities, population density, initial income and employment, previous population and employment growth, median age, proximity to metropolitan areas, and industry mix employment growth—the last a proxy for the local demand shock derived from the combination of initial county-level industry mix with national-level industry trends.

Hammond and Thompson (2008) and Glaeser et al. (1995) emphasize the contribution of education attainment to subsequent county and metropolitan area growth. We utilize a measure of education attainment based on obtaining a bachelor's degree (or higher) as well as two measures of human capital based on occupation: the share of workers in high human capital occupations and in occupations in the arts. The classifications follow the USDA's Economic Research Service definitions of creative class occupations and its artistic subset, respectively.[3] We treat the two as separate shares, so that the artistic occupations—those classified as "art and design workers" as well as "entertainers ... and related workers"—are excluded when calculating the high human capital share of employment.

Rappaport (2004, 2007) and Beeson et al. (2001) suggest persistent population growth in regions with higher amenities, which is supportive of employment growth. The amenity variable is taken from the USDA's Economic Research Service and reflects January and July temperatures, humidity, sunshine, topography and water coverage. Variables reflecting proximity to metropolitan areas were developed following Stephens and Partridge (2011). These variables reflect the propensity of counties to benefit from the spillover of urbanization economies generated by metropolitan areas of different sizes. The growth benefits of all distance variables are expected to decline with distance.

Partridge et al. (2009) include an industry mix employment growth variable as a measure of the local demand shock, a key control. For each industry j, we calculate the national growth rate ($G^j_{200--2007}$) for the period 2000–2007, and for each county i, we calculate the initial employment share in each industry (s_{ij}). The demand shock ($D^i_{2000--2007}$) is the sum of the products of national growth and initial local share: $D^i_{2000--2007} = \sum_j G^j_{2000--2007} * s^{ji}_{2000}$. Industry employment totals are at the three-digit NAICS level, except for farming and government employment, which are taken at the broad sector level. For suppressed values, we implement a simplified version of the technique outlined in Isserman and Westervelt (2006). Given initial employment, this local demand shock reflects the employment that would follow had each local industry behaved in line with its national counterpart. The variable will thus capture any trends related to the prominence of industries at the local level that are experiencing nation-wide growth or decline. Our entrepreneurship variables are left to explain the residual of local growth after controlling for the intersection of national trends with local industrial composition.

Our use of establishments birth and death data for entrepreneurial projects in Equation (4) warrants a brief discussion. We draw the establishment data from the Business Information Tracking Series. The birth and death rates are reported in units of births (or deaths) per thousand employees for the period 1998–1999, which is the earliest available period. In some regressions, we include the product of local establishment births and deaths to enable the measurement of dual effects from establishment deaths: the direct negative closure effect and the indirect positive information effect. When paired with the birth rate, the simple death rate includes both the negative closure effect and the positive information effect. Alone, the product summarizes the state of local entrepreneurial dynamism by effectively synthesizing overall entrepreneurial activity that generates useable market information; a high product and considerable information flow is possible even with a small net birth rate. When the birth rate, the death rate, and their product are combined, we are able to interpret the two opposing impacts of the death rate individually, with the death rate capturing the closure effect while the product isolates the information effect. A positive regression coefficient for the product variable provides a channel by which an additional death may have a measurably positive information effect capable of partially or even wholly offsetting the inarguable negative closure effect.[4]

---

[3]We prefer "high human capital" to "creative class" as the latter seems to imply a Bohemian character while the actual data includes accountants, managers, and lawyers.

While we view the novel entrepreneurial variables as an improved metric, their novelty argues for external confirmation. Three more common measures of entrepreneurship are thus included as well: the share of proprietors in total employment, the ten-year growth in the count of proprietors weighted by initial employment, and the estimated share of employment in firms with four or fewer employees. We include these measures for contiguity with other research and as a check of our hypotheses while noting potential drawbacks: proprietorship is a legal definition and the choice to operate as a corporation may reflect tax policies rather than intrinsic entrepreneurship, while a business's small size may simply reflect a lack of desire or ability to expand. We include these variables first alone, then each individually alongside the establishment birth rate, and finally altogether with the birth rate.

To alleviate concerns that we may misinterpret a non-causal relationship, we use an instrumental variable approach. Our instrumental approach utilizes historical mining employment data taken from the 1974 County Business Patterns. Our key instrument is the log of the estimated employment in the mining industry. Although mining employment shares may seem more appealing as instruments to reveal structures antithetical to entrepreneurship, the discrete existence of a non-trivial mining sector in fact may be more important in framing critical local market transactions. In particular, the higher wages and land rents inherent in such heavily-embedded mining operations create barriers to entry in the form of fundamental benchmarks for local factor markets, shifting the decision locus of both workers and firms towards such opportunities but away from the development of an entrepreneurial dynamic (Weiler, 1997, 2000b, 2001). Potential entrepreneurs are simultaneously burdened by these high costs while receiving few benefits from the large scale (and largely exported) mining activity (Graves et al., 2009). Within the spatial equilibrium model outlined previously, mining raises wages and land rents for all firms but raises productivity (and lowers input costs) only for those industries that either utilize the mines or directly benefit from their output—for instance, steel manufacturing.

To this end, we also use two non-linear instruments based on mining employment. First, we deploy indicator variables for counties with mining sector employment greater than different thresholds, which enables identification of areas where mining activity is likely to be export-oriented and thus pressure local factor prices. Second, we leverage the interaction between log mining employment and population density. While each of these instruments produces similar results when used individually, their use together produces the clearest perspective and so we focus our attention on the multiple-instrument regressions rather than on the single-instrument regressions, or on alternatives like the share of employment in the mining industry.[5] Finally, the period 1974 is chosen as it both has readily-available mining data and is sufficiently distant to be plausibly exogenous to employment growth in the post-2000 period, making it a potentially valid instrument.

One concern with the instrument may be that it proxies for locations in structural decline for non-entrepreneurial reasons. To counter this possibility, we bolster the case for the validity

---

[4]There is a high degree of collinearity between their product and births and deaths themselves. To account for this in the regression analysis, the interaction is the product of de-meaned births and deaths. The interpretation of marginal effects remains the same.
[5]To accommodate multiple potentially endogenous variables, in some instances we also utilize deep lags of alternative measures of entrepreneurship based on local proprietorship data.

of the instrumental approach by including the local demand shock and past employment growth as controls. In its construction, the first variable—which is the industry-mix growth term from shift-share analysis—directly utilizes 1998 local industry employment at the six-digit NAICS level, including mining industries. Therefore the regressions are analyzing solely the 2000–2007 "regional shift" from shift-share analysis. Furthermore, our inclusion of 1990–2000 employment growth as a control means that we are effectively testing for a relationship between the portion of the regional shift that is uncorrelated with local employment growth in the prior decade, further guarding against concerns that the instrument. If the lagged mining instruments were solely proxying for structural decline, the industry mix and lagged employment growth variables should render the instrument impotent. By accounting for industry structure in 1998, the first-stage relationship between our mining instruments and the residual local growth is more likely to be capturing the posited relationship

Due to the inclusion of these stringent controls—and in some regressions, the inclusion of state fixed effects—we see few alternative explanations for a relationship between historical mining activity and current employment growth besides our explanation: that historical mining activity dampened entrepreneurialism. As we shall show in the results section, multiple statistical tests confirm this hypothesis.

The main analysis focuses on employment growth from 2000–2007. The initial year is chosen for two reasons: first, it enables the use of control data from the 2000 decennial census, and second, it enables a peak-to-peak analysis across the business cycle. Non-census control variables are also from 2000. Before proceeding to these regressions, we explore the relationship between establishment births and deaths in 2000 and births in 2005. Our information hypothesis predicts that both births and deaths should be positively associated with future births. These regressions include the same control variables as the main regressions; the industry-mix growth term is adjusted to capture the period 2000–2005.

The primary unit of analysis is the county; in many instances we focus on the subsets of metropolitan or non-metropolitan counties. Counties offer the smallest-scale check of the hypotheses, given the lack of data available at smaller units. As counties vary in size from fewer than 100 residents to nearly ten million, we use initial-period employment to weight the observations for both the summary statistics and the regressions. We obtain similar results when using metropolitan areas as the unit of analysis; these results are reported in Appendix B.

The summary statistics for all variables are presented in Table 1. Additional detail about the data and variables used can be found in Appendix A.

4 Entrepreneurship and Information

We hypothesize that entrepreneurial projects create socially beneficial information about marketplace opportunities and thus induce faster economic growth. Through the failures and successes of entrepreneurial projects, local observers glean information that they are then

able to put to use in their own enterprises and funding decisions. In this sense, entrepreneurship breeds more entrepreneurship, promoting economic growth.

We first directly test our hypothesis by exploring whether our measures of entrepreneurial projects have a positive link to future establishment births. For this analysis, we utilize a similar framework to that presented above, and estimate a nearly identical regression to that implied by (3). In place of employment growth through 2007, the dependent variable is the establishment birth rate between 2005 and 2006.[6] Alongside the identical suite of control variables, we utilize the measures of entrepreneurial projects discussed above: the establishment birth rate, death rate, and their product. We expect to find a positive effect from all three variables as entrepreneurs and financiers are more willing to engage in risky entrepreneurial projects when those projects are made less uncertain by the information yielded from the success and failure of prior projects.

In Table 2, we present results that suggest this is the case. The first column suggests that a unit change in the birth rate yields a response in the latter-period birth rate of over 0.87—a very high degree of correlation, even in the presence of additional controls.[7] Alongside the establishment birth rate, we introduce the establishment death rate their product in the second column. Both the prior-period establishment death rate and the product of births and deaths have effects that are positive, statistically significant, and economically sizable: a one-point increase in the initial death rate is associated with an increase in the final birth rate of greater than one-fifth. For the interaction, the significance in the results of Column Two suggest that the magnitude of churning, whereby new enterprises replace or subsume older ones, has an independent positive effect independent of net establishment growth in the initial period. The birth and death rates remain significant in the presence of state fixed effects (Column Three) for all counties, and even for MSA fixed effects (Column Four) when examining the metropolitan subset. The significance of any churning effects on future births apparently become incorporated into state and MSA fixed effects, as indicated by the final two columns. Full results for the model are available in Table B.

5 Mining as an Instrumental Variable

For the employment growth regressions to follow, concerns about forward-looking entrepreneurs suggest the possibility of endogeneity. In the following section, we utilize historical mining activity as an instrumental variable. The logic, as described above, is that large concentrations of mining employment were likely to induce an economic ecosystem inimical to entrepreneurial activity. Mining activity will make for a useful instrument as, after including our stringent array of control variables, historical mining employment is likely to have little effect on employment growth—except through the entrepreneurship channel. However, its very strength as an instrument in the employment regressions makes the measure ill-suited as an instrument in the previous regressions. If (a lack of) mining activity induces a self-sustaining equilibrium with (high) low rates of entrepreneurial activity, then that mining activity cannot be thought of as exogenous to the dependent

---

[6]Similar results hold whether 2005 establishment birth rates are taken as total births divided by 1998, 2000, or 2005 employment. The alternatives ensure that the result is not mechanical; these results are available upon request.
[7]Complete results with control variable coefficients are available in the appendix for all regressions.

variable of future-period entrepreneurship. And, indeed, when mining is included in the regression alongside the three measures of entrepreneurial activity—births, deaths, and their product—the coefficient on mining is negative and significant at the 5 percent level and thus not exogenous.[8]

This finding suggests that mining activity is very strongly related to our chosen measure of entrepreneurship, suggesting that it may be a useful instrument in the growth regressions. To this effect, the first-stage results are presented in Table B in the appendix. These results, wherein we regress modern establishment birth rates on modern controls and the various instruments—including the key instruments of the log of 1974 mining employment—are strong and fit close to theory. In the main regression analyses, we use a suite of mining-related variables: the log of 1974 mining employment, nonlinear indicator variables for employment greater than 20 and 100 workers, an interaction between log of 1974 mining employment and local population density in 1980, and population density itself in 1980. The last two instruments follow from the hypothesis that smaller, less dense locales will have less latent entrepreneurial activity to lose, and that mining activity in dense counties will thus have a stronger negative impact; these coefficients have the predicted signs and are statistically significant in most specifications. For some regressions, we also make use of lagged versions of the alternative entrepreneurship measures based on proprietorship and small business employment.

The theoretical and empirical strength of the establishment birth and death results suggests that the combination of our novel measure of entrepreneurial information and the historical mining IV strategy offers a useful approach to analyzing the impact of entrepreneurship on economic growth. Furthermore, the uniformly strong results, including the positive relationships found between prior-year deaths and future establishment births, provide compelling initial evidence that an information effect is at work, whereby past failures significantly inform and foster future business creation. We utilize these findings and interpret the employment growth results in this light.

6 Results

The rate at which entrepreneurial projects are undertaken as measured by establishment births, deaths and their product is hypothesized to have a positive and significant effect on local employment growth. This is precisely what we find. We highlight the main specifications in Table 6 and include a range of alternative specifications to establish the robustness of the findings in the various appendix tables. Throughout the various permutations, the core result stands: that entrepreneurial projects have a significantly positive effect on employment growth. Finally, we incorporate the establishment birth rate, the establishment death rate, and their interaction to show that deaths have a measurable positive effect beyond the direct negative impact (as an establishment closure directly lowers employment). These results are shown in Table 6.

---

[8]These results are available from the authors upon request.

The first two columns of Table 6 present OLS results, the first column without and the second column with state fixed effects. The last two columns show the same results while using instrumental variables, the third column without state fixed effects and the fourth column including them. The results are remarkably consistent and consistently strong: for example, using the first column results, a one standard deviation increase in the establishment birth rate is associated with a 4.0 percentage-point increase in the growth rate of employment. With a mean employment growth rate of 8.75 percent, this is a meaningful movement. Furthermore, the stability of the coefficient to the introduction of state fixed effects in the second column suggests that the control variables are capturing a substantial portion of the variation across places and that the estimates are, in fact, capturing a relationship between entrepreneurial projects and employment growth.

One might be concerned that forward-looking entrepreneurs simply anticipate future growth and that these results suffer from reverse causality. These concerns are lessened by the similarly strong results from the third and fourth columns, which show IV results. First, the mining instruments are strong. Using the Cragg-Donald Wald F-statistic to test the strength of the instrument, the statistics are 47.5 and 45 for the two specifications. Second, the Hansen J-statistic is not significant, suggesting that the instruments are themselves exogenous to the estimating equation.[9] Finally, the GMM distance test fails to reject the null hypothesis that the establishment birth rate is exogenous (with P-values 0.55 and 0.17, respectively).[10] We thus focus the discussion on the OLS results.

To test the hypothesis that establishment births capture a unique and novel component of local entrepreneurialism, we have executed the same regressions as above while adding to the mix three traditional measures of entrepreneurship: the alternative measures are based on the share and growth rate of proprietorship and the share of employment in small businesses; these results can be found in Table B.[11] When including these measures as covariates, the birth rate results are virtually unchanged. Alone in OLS regressions, the alternatives show their expected positive signs, but the introduction of the establishment birth rate attenuates their effects. When using instruments for both the alternatives and for the establishment birth rate, no positive and significant relation- ship exists between any of the alternatives and employment growth. The birth rate, on the other hand, shows the expected positive relationship to employment growth in all cases. Entrepreneurial activity, as captured by the establishment birth rate, has a consistently strong positive effect on local economic growth.

As noted above, the covariates explain a significant portion of county-level variation in employment growth. By and large, these control variables show the expected relation- ships.

---

[9]In Table B in the appendix, we further test the exogeneity of the instruments by examining whether the inclusion of the establishment birth rate alongside the mining instruments reduces the (negative and significant) reduced-form relationship between the mining instruments and employment growth. In all cases, the mining instruments show no statistically significant relationship after the inclusion of the establishment birth rate—as expected.

[10]The GMM distance statistic tests whether the OLS and IV coefficient estimates are identical; rejection would have suggested that the IV estimates are appropriate.

[11]Results for further sensitivity tests are available upon request from the authors. These results consistently show a strong relationship between establishment births and employment growth with stable point estimates. In particular, the measure remains similar in magnitude even alongside the introduction of spatial dependence of both error terms and the dependent variable itself, as well as for the Census-designated metropolitan and nonmetropolitan subsets of counties. Nor does the period of measure for employment growth alter the results; using growth from 2000–2011 produces similar results.

For predicted employment growth, we find the expected positive and statistically significant relationship with observed employment growth in all cases. Workforce characteristics also significantly influence local job growth rates. In most cases, there is a negative relationship between the share of the workforce in arts occupations—entertainers and art and design workers—and employment growth.[12] A positive and statistically significant relationship is identified for high human capital occupations as classified by the USDA and including such occupations as managers, attorneys, and accountants. There is also an interesting interplay between the alternative methods of measuring human capital in counties: by occupation (as with HC Share) or by education (as with BA Share).

Employment growth was slower in counties with higher initial nominal incomes, consistent with the findings of Deller et al. (2001) for rural counties and Glaeser et al. (1995) for cities. No relationship was identified between initial employment levels and subsequent employment growth, but there was persistence in employment growth across time with positive and significant coefficient estimates on the 1990–2000 employment growth variable. When looking across the broad set of results included in the appendix, the census tract-weighted population density generally has a positive and statistically significant relationship, although this finding is reversed for non-metropolitan counties. Results also indicate slower employment growth in counties with a higher initial age. This result may reflect less potential for labor force growth among counties with an older population, although we note that Stephens and Partridge (2011) did not find a relationship between average age and employment in Appalachian and adjacent counties. Conditional on these other variables, no robust relationship was identified between amenities and employment growth.

One of the core predictions of the theoretical framework is that the termination of an entrepreneurial project provides information as to the local economic environment, and this information can be used by future agents in ways that should have a positive effect on economic growth. We have seen in the previous section that establishment deaths—our empirical approximation of project termination—do indeed have a positive effect on the future rate of entrepreneurial activity, and that, still further, the product of births and deaths itself has a positive effect. In Table 6, we explore whether the information derived from establishment deaths has measurable implications for subsequent employment growth.

Despite the handicap that establishment deaths have a direct negative effect on local employment, we find tentative evidence of the information-driven positive effect of establishment closures on eventual job growth, at least in metropolitan counties. The product of births and deaths has a positive coefficient for the metropolitan county subset and when using instruments—although the insignificance of establishment births in the instrumental variable regressions clouds interpretation. In the instrumental variable regressions, the three potentially endogenous variables are jointly significant and, for the regression with state fixed effects, we can reject at the 10 percent level that the instruments are weak and we fail to reject that the entrepreneurship measures are exogenous. We therefore focus again on the OLS regressions.

---

[12]Intriguingly, Table B reveals that the arts occupations' share of total employment may be positively related to the establishment birth rate—suggesting a possible relationship between so-called "creative classes" and entrepreneurial activity.

For metropolitan counties, the positive interaction provides a channel whereby a marginal establishment death increases the employment growth rate. That the interaction's positive effect shows up most clearly in the subset of metropolitan counties is not surprising: large metropolitan counties, with their deep labor markets, provide insurance against idiosyncratic job losses. Employees at closing metropolitan establishments can more easily find a job, enabling the informational effects of such closures to be identified more readily.

Of course, the direct effect of those establishment deaths is negative; a closure is still a closure. But in this model, the marginal effect of an establishment death in a particular county is the sum of the death rate coefficient and the product of that county's establishment birth rate with the "B*D" coefficient. For metropolitan counties, this comes to $-3.2 + (B_i - 4.34)$ *0.71 where B_i is the local establishment birth rate and 4.34 is the mean of metropolitan county birth rates. This implies that for a metropolitan county with a birth rate above 8.8, the marginal death has a *net positive* effect on future employment growth. Despite the different coefficient estimates, the comparable figure for the IV point estimates is 10.4. Although the closure job loss effect isn't fully offset for the large majority of counties that are below this threshold, even in these counties each incremental establishment death carries with it a smaller negative employment bite.

This finding runs contrary to the intuition that knock-on effects from the loss of one establishment might have negative repercussions for others; the reverse is in fact the case. Furthermore, as deaths are themselves highly correlated with births, all else is not equal: a marginal death carries with it a substantial fraction of an offsetting birth, further inducing employment growth. Along the lines of the theoretical framework outlined above, these *highly dynamic economies* feature an informational effect of establishment deaths that partially or fully offsets the negative direct effect.

The information content of deaths also helps to shape subsequent births, and the flip side of the noted result reinforces this implication of the information hypothesis. For the subset metropolitan of counties, the marginal effect of an establishment birth is estimated to be $5.2 + (D_i - 3.93)$ *0.71 where D_i is the establishment death rate and 3.93 is the mean death rate for metropolitan counties. A marginal birth thus has a larger effect on employment growth in the presence of more establishment deaths.

The information hypothesis again provides a framework for interpreting this result. Entrepreneurs and financiers in counties with many establishment deaths are able to draw on a deeper pool of failures as they shape new projects, sharpening their plans to avoid the pitfalls highlighted by previous failures and enhancing their projects' growth potential. The fact that new projects lead to more jobs in the presence of more establishment deaths underscores the significance of business failures to the growth process, a paradox which itself highlights both the role and importance of information flows.

While not dispositive, these twin novel findings of the indirectly positive impact of firm deaths fit comfortably within the marketplace information framework, dovetail with the results from the preliminary bridging results on the informational role of past entrepreneurship, and support the interpretation that entrepreneurial projects themselves

have a causal effect on economic growth. The apparent propensity of entrepreneurial activity to sustain itself suggests that localities are subject to *geographic information asymmetries* that manifest both within a place's level of entrepreneurship but also, ultimately, in its longer-term level of economic development.

7 Conclusion

This paper tests the proposition that entrepreneurial projects have a positive and determinative effect on employment growth in local economies. We introduce a novel set of entrepreneurial measures tailored to the information hypothesis: establishment births, establishment deaths, and their product. Using an instrumental variable approach ensures that causation is well-established, and after controlling for a variety of structural factors, the results indicate that entrepreneurial projects are indeed a causal determinant of future growth. This result holds for both counties and metropolitan areas and while using different control variables, different time periods, state fixed effects, models with spatial dependence, and alternative measures of entrepreneurship. Consistent with the information hypothesis, the strength of the relationship between establishment births and employment growth in the presence of alternative measures suggests that the relationship captures a unique component of entrepreneurialism.

We also provide evidence that current establishment births feed off past entrepreneurial projects: the establishment birth and death rates, as well as their product, enter positively in determining future rates of establishment births—even in the presence of a full suite of control variables. These results are entirely consistent with the entrepreneurial information framework outlined in the paper. This framework, in which future entrepreneurs, financiers, and existing firms draw information from the successes and failures of entrepreneurial projects, provides a consistent lens to understand the surprising yet revealing positive results from establishment deaths. The latter finding suggests that future participants utilize the richer information set generated by these failures to strengthen and accelerate their own eventual entrepreneurial innovations. Regional variations in such information sets leads to *geographical informational asymmetries*, which themselves create reinforcing cycles of business (in)activity and consequent economic growth (stagnation).

If entrepreneurial projects sustain employment growth and are themselves partially sustained by the information generated by entrepreneurial failures, then those past failures ought to have a positive and measurable effect on employment growth. And, within metropolitan counties, firm deaths do have indirect positive effects on employment growth. These positive effects flow through two channels. First, information from a broader set of past failures will allow the marginal firm birth to have greater job-creation effects. Second, the most *highly dynamic economies* with sufficiently high establishment birth rates may see faster employment growth from the marginal establishment death, while establishment closures in less dynamic economies produce a positive informational effect that partially offsets the negative direct effects. These twin novel findings—alongside the instrumental variable and other analyses—provides powerful support to the notion that information from entrepreneurial projects is an important causal input to local economic growth.

These results underline the importance of entrepreneurship in economic development and push the limits of the current conception of entrepreneurship. Understanding entrepreneurship as a process of project-based information revelation—in addition to the well-known role it serves as a conduit for innovation—encourages both policymakers and researchers to see entrepreneurship anew, and to incorporate the unique and crucial actions of entrepreneurs into both academic research and practical policy.

Acknowledgments

References

Àcs, Z.; Armington, C. Entrepreneurship, Geography, and American Economic Growth. Cambridge University Press; 2006.

Àcs Z, Braunerhjelm P, Audretsch D, Carlsson B. The knowledge spillover theory of entrepreneurship. Small Business Economics. 2009; 32(1):15–30.

Akerlof G. The market for 'lemons': Quality uncertainty and the market mechanism. The Quarterly Journal of Economics. 1970:488–500.

Audretsch D, Keilbach M. Entrepreneurship, Capital and Economic Performance. Regional Studies. 2004; 38(8):949–959.

Audretsch D, Keilbach M. The localisation of entrepreneurship capital: Evidence from germany. Papers in Regional Science. 2007; 86(3):351–365.

Basu N, Chevrier M. Distance, information asymmetry, and mergers: Evidence from canadian firms. Managerial Finance. 2011; 37(1):21–33.

Beeson P, DeJong D, Troesken W. Population growth in us counties, 1840–1990. Regional Science and Urban Economics. 2001; 31(6):669–699.

Blanchard O, Katz L, Hall R, Eichengreen B. Regional evolutions. Brookings papers on economic activity. 1992; 1992(1):1–75.

Braunerhjelm P, Àcs ZJ, Audretsch DB, Carlsson B. The missing link: Knowledge diffusiong and entrepreneurship in endogenous growth. Small Business Economics. 2010; 34(2):105–125.

Chinitz B. Contrasts in Agglomeration: New York and Pittsburgh. American Economic Review. 1961; 51(2):279–289.

Deller S, Tsai T, Marcouiller D, English D. The role of amenities and quality of life in rural economic growth. American Journal of Agricultural Economics. 2001; 83(2):352–365.

Glaeser E, Kerr S, Kerr W. Entrepreneurship and urban growth: An empirical assessment with historical mines. Working. 2012

Glaeser E, Scheinkman J, Shleifer A. Economic growth in a cross-section of cities. Journal of Monetary Economics. 1995; 36(1):117–143.

Graves P, Weiler S, Tynon E. The economics of ghost towns. Journal of Regional Analysis and Policy. 2009; 39(2):131–140.

Hall R, Gordon R, Holt C. Why is the unemployment rate so high at full employment? Brookings Papers on Economic Activity. 1970; 1970(3):369–410.

Hammond G, Thompson E. Determinants of income growth in metropolitan and nonmetropolitan labor markets. American Journal of Agricultural Economics. 2008; 90(3):783–793.

Huang T, Orazem P, Wohlgemuth D. Rural population growth, 1950–1990: the roles of human capital, industry structure, and government policy. American Journal of Agricultural Economics. 2002; 84(3):615–627.

Isserman AM, Westervelt J. 1.5 million missing numbers: Overcoming employment suppression in county business patterns data. International Regional Science Review. 2006; 29(3):311–335.

Jacobs, J. The Economy of Cities. Vintage; New York: 1970.

Krugman P. The narrow moving band, the dutch disease, and the competitive consequences of mrs. thatcher: Notes on trade in the presence of dynamic scale economies. Journal of development Economics. 1987; 27(1):41–55.

Krugman P. Increasing Returns and Economic Geography. The Journal of Political Economy. 1991; 99(3):483–499.

Lang W, Nakamura L. A model of redlining. Journal of Urban Economics. 1993; 33(2):223–234.

Low SA, Weiler S. Employment risk, returns, and entrepreneurship. Economic Development Quarterly. 2012; 26(3):238–251.

Lucas R. On the mechanics of economic development. Journal of monetary economics. 1988; 22(1):3–42.

Moretti E. Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics. 2004; 121:175–212.

Moretti E. Local labor markets. Handbook of Labor Economics. 2011; 4:1237–1313.

Partridge M, Rickman D, Ali K, Olfert R. Do new economic geography agglomeration shadows underlie current population dynamics across the urban hierarchy. Papers in Regional Science. 2009; 88(2):445–466.

Rappaport J. Why are population flows so persistent? Journal of Urban Economics. 2004; 56(3):554–580.

Rappaport J. Moving to nice weather. Regional Science and Urban Economics. 2007; 37(3):375–398.

Roback J. Wages, Rents, and the Quality of Life. The Journal of Political Economy. 1982; 90(6):1257–12. T8.

Romer P. Increasing returns and long-run growth. The Journal of Political Economy. 1986:1002–1037.

Romer P. Endogenous technological change. Journal of Political Economy. 1990; 98(5)

Shapiro J. Smart cities: Quality of life, productivity, and the growth effects of human capital. The Review of Economics and Statistics. 2006; 88(2):324–335.

Stephens H, Partridge M. Do entrepreneurs enhance economic growth in lagging regions? Growth and Change. 2011; 42(4):431–465.

Stephens HM, Partridge MD, Faggian A. Innovation, entrepreneurship and economic growth in lagging regions. Journal of Regional Science. 2013

Summers LH, Abraham KG, Wachter ML. Why is the unemployment rate so very high near full employment? Brookings Papers on Economic Activity. 1986; 1986(2):339–396.

Uysal V, Kedia S, Panchapagesan V. Geography and acquirer returns. Journal of Financial Intermediation. 2008; 17(2):256–275.

Weiler S. The economics of the struggling structurally unemployed. Journal of Appalachian Studies. 1997; 3(1):71–98.

Weiler S. Pioneers and settlers in lo-do denver: Private risk and public benefits in urban redevelopment. Urban Studies. 2000a; 37(1):167–179.

Weiler S. Unemployment in regional labor markets: Using structural theories to understand local jobless rates in west virginia. Industrial and Labor Relations Review. 2000b; 54:573.

Weiler S. Unemployment in regional labor markets: Using structural theories to understand local jobless rates in west virginia. Industrial and Labor Relations Review. 2001:573–592.

Weiler S. A park by any other name: National park designation as a natural experiment in signaling. Journal of Urban Economics. 2006; 60(1):96–106.

Weiler S, Hoag D, Fan C. Prospecting for economic returns to research: Adding informational value at the market fringe. Journal of Regional Science. 2006; 46(2):289–311.

A Data

Data on establishment dynamics comes from the Census Bureau's Business Information Tracking Series. Establishment births and deaths are divided by population data from the *Regional Economic Information System* to yield birth and death rates. In some regressions, we also include the de-meaned product of the birth and death rates. We utilize 1998–1999 birth and death data as the base period as it is the earliest with readily-available data.

The instrumental variables feature data drawn from the 1974 *County Business Patterns* from the Bureau of Census. The mining employment data was, in some cases, suppressed and only a range was reported. In these instances, total employment was estimated using the midpoint of the smallest range possible based on, first, the reported range for county mining employment and, second, the range implied by the counts of establishments in different employment size bins. A county with a reported mining employment range of 10–19 and two mining establishments each with employment in the range of 6–10 would have a narrowed range of 12–19, and an estimated employment of 16.5. When calculating the log of mining employment, the final employment counts and estimates were increased by one employee to accommodate counties with no employment.

The other measures of entrepreneurship utilized in the robustness checks are based on proprietorship and the share of employment in businesses with fewer than five employees. The two proprietorship measures are the share of proprietors in total employment in 1998 and the growth rate of proprietors from 1988 to 1998 weighted by 1988 employment. Data on proprietors is from the *Regional Economic Information System* of the Bureau of Census, as is the total employment data. The share of employment in small businesses with 1 to 4 employees during 1998 was estimated using *County Business Patterns* data from the Bureau of Census. In IV results using these measures, lagged values of these variables from, respectively, 1969, the period 1969–1979, and 1974 were used as instruments alongside the mining and density variables.

The key dependent variable is the rate of growth in non-farm employment between 2000 and 2007, a peak-to-peak period in terms of the business cycle. Robustness results include growth from 2000–2011. Data on the growth variables for the relevant years come from the *Regional Economic Information System* produced by the Bureau of Economic Analysis of the U.S. Department of Commerce.

The data for occupations in the arts and requiring a high degree of human capital are taken from the USDA's Economic Research Service as explained in the text. The population and the share of the population with a bachelor's degree (including those with additional degrees) is taken from the 2000 Census.

Tract-weighted population density for the year 2000 is constructed using tract-level data on population and land area. Tract population density is weighted by tract population and summed to the county level. The tract-weighted measure is included as traditional population density measures may be heavily influenced by the non-urban area of the county, which is quite heterogeneous: Los Angeles County is much larger than the city and contains

portions of multiple national forests, whereas Manhattan is its own county. The key results are similar when using either measure.

Log income in 2000, log employment in 2000, and employment growth from 1990–2000 are constructed using REIS data. The local demand shocks for 2000–2007 (for most regressions) and 2000–2011 (for robustness regressions) use a combination of REIS data for government employment with CBP data for most industries. The local demand shock variables are constructed using employment totals for primarily three-digit NAICS codes, as described in the text. For suppressed values, we implement a simplified version of the technique outlined in Isserman and Westervelt (2006) that is similar to that described above for the mining instruments.

Median age in 2000 is taken from Census Bureau estimates. Population growth in the period 1950–1960 is taken from the decennial censuses. The amenity scores are the standardized scores generated by Economic Research Service of the United States Department of Agriculture. For the distance measures, MSA populations and distances are taken from and constructed using county populations in the 2000 Census.

The analysis in the paper focuses on two geographic levels: counties (and county-equivalents) and metropolitan statistical areas. The possibility of geographic information asymmetries encourages a focus on local areas that are likely to contain the informational spillovers generated by entrepreneurial or research activity. Counties are the smallest geographic unit for which the relevant data—notably on establishment births and deaths, but other control variables as well—is readily available, and so we focus our analysis there. Metropolitan areas are also included as inter-county commuting ties suggest that information may flow across these jurisdictional boundaries.

The United States is divided into 3,143 counties and county equivalents, including a number of independent cities that are economically integrated with their surrounding counties. Upon aggregation of independent cities with surrounding counties, 3,111 counties remain. Of these, we focus on the lower forty-eight states due to concerns about data and applicability. In addition, there are a number of instances of county formations during the period in question; these counties are aggregated to their largest extent in order to ensure data consistency. This leaves a sample of 3,072 counties, county equivalents, and county aggregates that encompass the entirety of the continental United States. When aggregating these finalized counties to the relevant metropolitan areas according to year 2000 definitions, we end up with 356 metropolitan areas.

B Tables

Table 5

Preliminary Entrepreneurship Results. Dependent variable is the establishment birth rate, 2005. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 1 and the data appendix for details on the variables.

Birth Rate, 1998	0.87*** (0.025)	0.7*** (0.038)	0.68*** (0.036)	0.62*** (0.09)
Death Rate, 1998	—	0.23*** (0.039)	0.23*** (0.04)	0.35*** (0.1)
B*D, 1998	—	0.036** (0.015)	0.01 (0.014)	0.038 (0.026)
Demand Shock, 2000–07	0.007 (0.007)	−0.001 (0.006)	0.004 (0.006)	−0.001 (0.019)
Emp Growth, 1990–2000	0.007*** (0.002)	0.009*** (0.002)	0.01*** (0.002)	0.012*** (0.003)
Log Employment, 2000	0.11*** (0.029)	0.11*** (0.029)	0.09*** (0.025)	0.05 (0.05)
Log Income, 2000	−0.79*** (0.162)	−0.79*** (0.159)	−0.57*** (0.14)	−0.3 (0.3)
Density, 2000	0.004 (0.004)	0.003 (0.003)	−0.004 (0.003)	−0.001 (0.005)
HC Share, 2000	0.030** (0.013)	0.032*** (0.012)	0.04*** (0.012)	0.022 (0.028)
Arts Share, 2000	−0.028 (0.078)	−0.026 (0.076)	0.035 (0.07)	−0.07 (0.12)
BA Share, 2000	−0.012 (0.008)	−0.011 (0.007)	−0.024*** (0.007)	−0.016 (0.018)
Pop Growth, 1950–1960	0.047 (0.060)	0.017 (0.07)	−0.1 (0.06)	−0.23*** (0.07)
Median Age, 2000	0.016** (0.007)	0.007 (0.007)	−0.005 (0.008)	−0.005 (0.019)
Amenity Score	0.043*** (0.008)	0.042*** (0.008)	0.024* (0.012)	−0.034 (0.031)
Distance to MSA	0.002*** (0.001)	0.003*** (0.001)	0.002*** (0.001)	0.009 (0.007)
Marginal distance to MSA > 250k	−0.001 (0.0001)	0 (0.001)	0 (0.000)	−0.009 (0.008)
Marginal distance to MSA > 500k	0.002*** (0.001)	0.002*** (0.001)	0 (0.000)	−0.001 (0.006)
Marginal distance to MSA >1M	−0.001*** (0.000)	−0.001*** (0.000)	−0.001** (0.000)	0.005 (0.004)
Constant	6.66*** (1.55)	6.69*** (1.54)	4.89*** (1.37)	3.81 (3.44)
N	3072	3072	3072	825
R^2	0.78	0.79	0.82	0.95
F (non-FEs)	272	242	89.9	27.5
Counties	All	All	All	Metro
Fixed Effects	No	No	State	MSA
Full Controls	Yes	Yes	Yes	Yes
Regression	OLS	OLS	OLS	OLS

Table 6

First Stage Results. Dependent variable is the establishment birth rate, 1998. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 1 and the data appendix for details on the variables.

Log Mining Emp, 1974	−0.08*** (0.022)	0.016 (0.037)	0.034 (0.034)	−0.1 (0.06)
Mining Emp > 20	—	−0.12 (0.12)	−0.14 (0.1)	−0.1 (0.2)

Mining Emp > 100	—	−0.24** (0.11)	−0.12 (0.1)	−0.2 (0.22)
Density, 1980	—	0.1 (0.07)	0.09 (0.07)	0.31** (0.14)
Lagged Mining×Density	—	−0.011*** (0.003)	−0.011*** (0.003)	−0.02*** (0.006)
Predicted Employment Growth, 2000–07	0.08*** (0.009)	0.07*** (0.009)	0.07*** (0.009)	0.05*** (0.018)
Lagged Employment Growth, 1990–2000	0.014*** (0.003)	0.014*** (0.003)	0.003 (0.002)	0.008* (0.005)
Log Employment, 2000	0.11* (0.05)	0.035 (0.05)	−0.043 (0.042)	0.022 (0.09)
Log Income, 2000	−1.8*** (0.33)	−1.61*** (0.28)	−1.11*** (0.27)	−0.28 (0.54)
Density, 2000	0.01 (0.007)	−0.005 (0.029)	0.000 (0.03)	−0.09 (0.06)
HC Share, 2000	0.08*** (0.022)	0.09*** (0.022)	0.08*** (0.02)	−0.06 (0.046)
Arts Share, 2000	0.047 (0.12)	0.28** (0.12)	0.27** (0.11)	0.55** (0.25)
BA Share, 2000	−0.03*** (0.014)	−0.037*** (0.014)	−0.045*** (0.012)	−0.015 (0.025)
Pop Growth, 1950–1960	0.26** (0.13)	0.3* (0.14)	0.29* (0.15)	0.6 (0.54)
Median Age, 2000	0.08*** (0.011)	0.08*** (0.011)	0.08*** (0.009)	0.11*** (0.023)
Amenity Score	0.1*** (0.012)	0.08*** (0.014)	0.07*** (0.014)	0.012 (0.031)
Distance to MSA	0.004*** (0.001)	0.004*** (0.001)	0.002*** (0.001)	0.001 (0.001)
Marginal distance to MSA > 250k	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	−0.002** (0.001)
Marginal distance to MSA > 500k	0.001* (0.001)	0.001* (0.001)	0.001 (0.000)	0.002** (0.001)
Marginal distance to MSA > 1M	−0.001*** (0.000)	−0.001** (0.000)	0.000 (0.000)	−0.001 (0.001)
Constant	17.2*** (3.11)	15.4*** (2.7)	11.3*** (2.53)	0.19 (5.15)
N	3072	3072	3072	3072
R^2	0.49	0.52	0.44	0.16
F—Regression	42.1	37.6	37.9	5.7
F—IV	13.9	8.87	8.26	6.86
Dependent Variable	Births	Births	Deaths	Births×Deaths
Counties	All	All	All	All
Fixed Effects	No	No	No	No
Regression	OLS	OLS	OLS	OLS

Table 7

Robustness. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 1 and the data appendix for details on the variables. The weak instrument-robust p-value is for the Anderson-Wald F test statistic; the weak instrument test is the Cragg-Donald Wald F statistic. The test for the endogeneity of entrepreneurship (i.e., the birth rate) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J statistic.

Birth Rate, 1998	—	3.79*** (0.36)	—	3.32* (1.72)
Proprietorship Share, 1998	−0.12 (0.14)	−0.13 (0.13)	−0.14 (0.32)	−0.6** (0.3)
Proprietorship Growth Rate, 1988–98	0.63*** (0.17)	0.67*** (0.14)	4.02 (2.56)	1.4 (1.49)

Small Business Employment Share, 1998	0.7*** (0.19)	−0.31 (0.19)	−1.46 (1.18)	−0.13 (0.81)
Predicted Employment Growth, 2000–07	0.81*** (0.12)	0.69*** (0.11)	1.29*** (0.36)	0.76** (0.32)
Lagged Employment Growth, 1990–2000	0.18*** (0.047)	0.12*** (0.038)	−0.27 (0.35)	0.043 (0.18)
Log Employment, 2000	0.13 (0.45)	−0.33 (0.42)	−0.17 (0.76)	−0.26 (0.44)
Log Income, 2000	−13.3*** (3.94)	−10.1*** (3.5)	0.98 (12.7)	−10.8 (8.14)
Density, 2000	0.09** (0.037)	0.08** (0.035)	0.04 (0.08)	0.06 (0.045)
HC Share, 2000	0.91*** (0.27)	0.67*** (0.25)	0.3 (0.55)	0.74** (0.36)
Arts Share, 2000	−1.78* (0.93)	−1.96** (0.86)	−2.62** (1.33)	−2.07** (0.91)
BA Share, 2000	−0.42*** (0.14)	−0.33** (0.13)	−0.14 (0.27)	−0.34* (0.18)
Population Growth, 1950–1960	1.99** (0.91)	0.93 (0.78)	4.18* (2.23)	1.28 (1.72)
Median Age, 2000	−0.35*** (0.11)	−0.52*** (0.1)	−0.23 (0.16)	−0.44*** (0.14)
Amenity Score	0.2 (0.15)	0.016 (0.14)	−0.49 (0.66)	0.09 (0.37)
Distance to MSA	−0.017* (0.009)	−0.018** (0.008)	−0.019 (0.015)	−0.017* (0.009)
Marginal distance to MSA > 250k	−0.004 (0.004)	−0.001 (0.004)	0.006 (0.009)	0.003 (0.005)
Marginal distance to MSA > 500k	−0.002 (0.004)	−0.004 (0.004)	−0.003 (0.006)	−0.004 (0.004)
Marginal distance to MSA > 1M	0.001 (0.004)	0.004 (0.003)	0.009 (0.008)	0.003 (0.005)
Constant	136*** (38)	109*** (33.4)	−4.54 (127)	118 (81.7)
N	3072	3072	3072	3072
R^2	0.53	0.58	—	0.56
F	49.1	55.6	17.6	45.4
Counties	All	All	All	All
Fixed Effects	No	No	No	No
Regression	OLS	OLS	IV	IV
Weak ID F-test			4.72	1.25
Weak IV-robust p-value	—	—	0.5	0.046
Endogeneity of Entrepreneurship p-value	—	—	0.001	0.005
Endogeneity of IVs p-value		—	—	0.93

Table 8

Full Entrepreneurial Information Results. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 1 and the data appendix for details on the variables. The weak instrument-robust p-value is for the Anderson-Wald F test statistic; the weak instrument test is the Cragg-Donald Wald F statistic. The test for the endogeneity of entrepreneurship (i.e., the birth rate) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J statistic.

Birth Rate, 1998	4.73*** (0.56)	5.2*** (0.75)	5.37*** (0.75)	7.74* (4.32)	7.31* (4.35)
Death Rate, 1998	−1.84*** (0.5)	−3.2*** (0.84)	−2.97*** (0.84)	−9.44* (5.07)	−6.11 (4.64)
Births×Deaths, 1998	0.17 (0.19)	0.71** (0.28)	0.49** (0.22)	1.99** (0.94)	1.05 (0.84)
Predicted Employment Growth, 2000–07	0.67*** (0.11)	0.83*** (0.18)	0.68*** (0.16)	0.93*** (0.18)	0.64*** (0.12)

Lagged Employment Growth, 1990–2000	0.19*** (0.038)	0.24*** (0.032)	0.26*** (0.032)	0.15** (0.07)	0.16** (0.07)
Log Employment, 2000	−0.31 (0.42)	−0.11 (0.49)	−0.08 (0.51)	−0.35 (0.49)	−0.15 (0.45)
Log Income, 2000	−12.5*** (3.39)	−14.7*** (4.01)	−16.2*** (3.51)	−16.1*** (4.66)	−15.2*** (4.13)
Density, 2000	0.1*** (0.035)	0.11** (0.042)	0.06 (0.04)	0.13*** (0.043)	0.06 (0.039)
HC Share, 2000	0.81*** (0.28)	0.63** (0.31)	0.45* (0.25)	1.26*** (0.31)	0.92*** (0.3)
Arts Share, 2000	−1.83** (0.89)	−1.22 (1.06)	0.54 (0.96)	−2.01** (1.02)	−0.43 (0.9)
BA Share, 2000	−0.41*** (0.14)	−0.35* (0.18)	−0.3** (0.14)	−0.67*** (0.17)	−0.52*** (0.14)
Pop Growth, 1950–1960	0.67 (0.84)	−0.48 (0.83)	−0.013 (0.77)	0.7 (0.99)	0.25 (0.85)
Median Age, 2000	−0.5*** (0.1)	−0.36*** (0.13)	−0.53*** (0.13)	−0.33** (0.16)	−0.62*** (0.12)
Amenity Score	0.15 (0.14)	0.17 (0.15)	−0.22 (0.27)	0.47** (0.18)	−0.16 (0.25)
Distance to MSA	−0.017* (0.009)	0.06*** (0.023)	−0.048** (0.022)	0.000 (0.005)	−0.005 (0.01)
Marginal distance to MSA >250k	−0.003 (0.004)	−0.008 (0.005)	−0.02** (0.008)	−0.004 (0.005)	−0.011* (0.006)
Marginal distance to MSA >500k	−0.004 (0.004)	0.000 (0.005)	0.005 (0.007)	0.005 (0.005)	−0.001 (0.005)
Marginal distance to MSA >1M	0.004 (0.003)	0.006 (0.004)	0.003 (0.005)	0.408 (0.422)	0.006 (0.005)
Constant	132*** (32.5)	150*** (37.8)	176*** (33.7)	175*** (45.3)	−173*** (40.2)
N	3072	825	825	3072	3072
R^2	0.56	0.65	0.74	0.44	0.6
F	59.8	50.6	30.7	41.5	35.6
Counties	All	Metro	Metro	All	All
Fixed Effects	No	No	No	No	Yes
Regression	OLS	OLS	OLS	IV	IV
Weak ID F-test				5.98	7.86
Weak IV-Robust P-value	—	—	—	0.046	0.13
Endogeneity of Entrepreneurship P-value	—	—	—	0.5	0.99
Endogeneity of IVs P-value				0.99	0.9

Table 1

Summary Statistics. Observations are weighted by year 2000 employment unless otherwise noted. Observations total 3,072 for all variables except the metropolitan county subset for which there are 825 observations. See the data appendix for more details.

Dependent Variables	Mean	Std Dev	Min	Max
Employment Growth Rate, 2000–2007	8.75	11.7	−35.3	169
Employment Growth Rate, 2000–2011	6.28	13.5	−38.2	214

Entrepreneurship Variables	Mean	Std Dev	Min	Max
Establishment Births, 1998–99	4.32	1.1	0	12.6
Establishment Deaths, 1998–99	3.95	0.9	0	17.2
Establishment Births, 1998–99 (Metro Counties)	4.34	1.04	0.49	10.4
Establishment Deaths, 1998–99 (Metro Counties)	3.93	0.84	0.43	9.08
Establishment Births, 2005	4.72	1.33	0	23.1
Establishment Deaths, 2005	4.05	1.06	0	28.9
Log of Mining Employment (plus one), 1974	5.2	2.27	0	9.98
Mining Employment Greater than 20, 1974	0.84	0.36	0	1
Mining Employment Greater than 100, 1974	0.62	0.48	0	1
Population Density, 1980	2.4	8.39	0	62.2
Share of Proprietors, 1998	15.4	4.81	1.74	56.9
Share of Proprietors, 1979	12.2	4.01	0.78	41.1
Employment-Weighted Proprietor Growth, 1988–98	4.62	4.77	−13.7	99.3
Employment-Weighted Proprietor Growth, 1969–79	6.07	6.65	−8.17	106
Employment Share in Small Establishments, 1998	5.15	2.06	0.65	100
Employment Share in Small Establishments, 1974	4.25	2.08	0	100

Control Variables	Mean	Std Dev	Min	Max
High Human Capital Share, 2000	24.6	7.14	0	53
Arts Share, 2000	1.25	0.81	0	6.67
Bachelor's Degree Share, 2000	26	9.85	4.92	60.5
Tract-Weighted Population Density, 2000	6.5	15.4	0	114
Log Income, 2000	10.6	0.28	8.03	11.4
Log Employment, 2000	12.4	1.68	4.55	15.5
Employment Growth, 1990–2000	22.2	20.6	−39.4	767
Predicted Employment Growth, 2000–2007	6.5	2.82	−16.5	86.4
Median Age, 2000	35.2	3.2	20.6	54.3
Population Growth, 1950–1960	0.39	0.48	−0.42	3.71
Amenity Score	1.3	3.35	−6.4	11.2
Distance to Nearest MSA in miles	26	31.8	0	371
Marginal Distance to MSA > 250,000	41.7	64.3	0	762
Marginal Distance to MSA > 500,000	58.3	81.5	0	797
Marginal Distance to MSA > 1,000,000	78.3	97.8	0	797

Table 2

Preliminary Entrepreneurship Results. Dependent variable is the establishment birth rate, 2005. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See the Table 1 and the data appendix for details on the variables. See Table B in the appendix for full results including for control variables.

Birth Rate, 1998	0.87*** (0.025)	0.7*** (0.038)	0.68*** (0.036)	0.62*** (0.09)
Death Rate, 1998	—	0.23*** (0.039)	0.23*** (0.04)	0.35*** (0.1)
Births × Deaths, 1998	—	0.036** (0.015)	0.01 (0.014)	0.038 (0.026)
N	3072	3072	3072	825
R^2	0.78	0.79	0.82	0.94
F-test for non-FEs variables	272	242	90	27.5
Counties	All	All	All	Metro
Fixed Effects	No	No	State	MSA
Full controls	Yes	Yes	Yes	Yes
Regression	OLS	OLS	OLS	OLS

Table 3

Main Results. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See Table 1 and the data appendix for details on the variables. The instruments used are as follows: the log of mining employment in 1974, indicator variables for mining employment greater than 20 and 100 in 1974, the population density in 1980, and the interaction between log of mining employment and population density. The weak instrument-robust P-value is for the Anderson-Wald F-test statistic; the weak instrument test is the Cragg-Donald Wald F-statistic. The test for the endogeneity of entrepreneurship (i.e., the birth rate) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J-statistic. See Table B for robustness tests using alternative definitions of entrepreneurship alongside establishment births. Further robustness tests are available from the authors upon request.

Birth Rate, 1998	3.62*** (0.39)	3.61*** (0.38)	3.24*** (1.04)	2.18** (0.97)
Predicted employment growth, 2000–07	0.61*** (0.11)	0.55*** (0.1)	0.64*** (0.13)	0.63*** (0.12)
Lagged Emp Growth, 1990–2000	0.2*** (0.038)	0.22*** (0.045)	0.21*** (0.041)	0.23*** (0.047)
Log Employment, 2000	−0.27 (0.43)	−0.22 (0.43)	−0.26 (0.44)	−0.21 (0.44)
Log Income, 2000	−12.2*** (3.43)	−14.1*** (3.18)	−12.9*** (4.37)	−16.8*** (4.37)
Density, 2000	0.1*** (0.035)	0.03 (0.029)	0.1** (0.04)	0.037** (0.031)
HC Share, 2000	0.75*** (0.28)	0.71*** (0.24)	0.78*** (0.28)	0.91*** (0.27)
Arts Share, 2000	−1.8** (0.9)	−0.12 (0.78)	1.81** (0.9)	−0.039** (0.77)
BA Share, 2000	−0.37*** (0.14)	−0.4*** (0.11)	−0.38** (0.14)	−0.377*** (0.12)
Pop Growth, 1950–1960	0.61 (0.83)	0.53 (0.72)	0.72 (0.91)	0.79 (0.76)
Median Age, 2000	−0.55*** (0.11)	−0.78*** (0.1)	−0.52*** (0.13)	−0.69*** (0.12)
Amenity Score	0.11 (0.14)	−0.27 (0.24)	0.14 (0.17)	−0.1 (0.26)
Distance to MSA	−0.017* (0.009)	−0.004 (0.009)	−0.02 (0.01)	0.001 (0.01)
Marg dist MSA > 250k	−0.004 (0.004)	−0.015*** (0.005)	−0.004 (0.004)	−0.015*** (0.006)
Marg dist MSA > 500k	−0.004 (0.004)	−0.002 (0.005)	−0.004 (0.004)	−0.001 (0.005)
Marg dist MSA > 1M	0.003 (0.003)	0.005 (0.004)	0.003 (0.004	0.002 (0.005)
Constant	129*** (32.8)	158*** (31.2)	136*** (42.4)	193*** (43.7)
N	3072	3072	3072	3072
R^2	0.55	0.64	0.55	0.63
F (non-FEs)	59.0	41.0	55.2	41.7
Counties	All	All	All	All
Fixed Effects	No	State	No	State
Full Controls	Yes	Yes	Yes	Yes
Regression	OLS	OLS	IV	IV
Weak IV-Robust F-test	—	—	47.5	45.0
Weak IV-Robust P-value			0.02	0.41
Endog of E'ship P-value	—	—	0.55	0.17
Endog of IVs P-value			0.25	0.99

Table 4

Results with Birth and Death Rates. Dependent variable is the employment growth rate, 2000–2007. Coefficients with one, two, and three stars are significant at the 10%, 5%, and 1% levels, respectively, using a two-tailed test. See Table 1 and the data appendix for details on the variables. The instruments used are as follows: the log of mining employment in 1974, indicator variables for mining employment greater than 20 and 100 in 1974, the population density in 1980, and the interaction between log of mining employment and population density, the lagged proprietor's share of employment in 1979, the employment-weighted growth rate of proprietorship from 1969–1979, and the share of employment in small businesses in 1974. For first-stage results, see Table B2. The weak instrument-robust P-value is for the Anderson-Wald F-test statistic; the weak instrument test is the Cragg-Donald Wald F-statistic. The test for the endogeneity of entrepreneurship (i.e., the birth and death rates) is the GMM distance measure; the test for the endogeneity of the instrument is the Hansen's J-statistic.

Birth Rate, 1998	4.73*** (0.56)	5.2*** (0.75)	5.37*** (0.75)	7.74* (4.32)	7.31* (4.35)
Death Rate, 1998	−1.84*** (0.5)	−3.2*** (0.84)	−2.97*** (0.84)	−9.44* (5.07)	−6.11 (4.64)
Births×deaths, 1998	0.17 (0.19)	0.71** (0.28)	0.49** (0.22)	1.99** (0.94)	1.05 (0.84)
N	3072	825	825	3072	3072
R^2	0.56	0.65	0.74	0.44	0.6
F	59.8	50.6	30.7	41.5	35.6
Counties	All	Metro	Metro	All	All
Fixed Effects	No	No	Yes	No	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes
Regression	OLS	OLS	OLS	IV	IV
Weak IV-Robust F-test	—	—	—	5.98	7.86
Weak IV-Robust P-value	—	—	—	0.046	0.13
Endogeneity of Entrepreneurship P-value	—	—	—	0.5	0.99
Endogeneity of IVs P-value	—	—	—	0.99	0.9