December 2002

Who are the Entrepreneurs? Evidence from Taxpayer Data

Donald Bruce
University of Tennessee

Douglas Holtz-Eakin
Syracuse University

Follow this and additional works at: https://digitalcommons.pepperdine.edu/jef

Recommended Citation
Available at: https://digitalcommons.pepperdine.edu/jef/vol7/iss1/2

This Article is brought to you for free and open access by the Graziadio School of Business and Management at Pepperdine Digital Commons. It has been accepted for inclusion in The Journal of Entrepreneurial Finance by an authorized editor of Pepperdine Digital Commons. For more information, please contact josias.bartram@pepperdine.edu, anna.speth@pepperdine.edu.
INTRODUCTION

The United States has experienced a revival of interest in entrepreneurs. While much of the public fascination has focused on “dot com” millionaires and similar high-profile phenomena, a coincident surge in research has focused on the economic foundations of this boom. In executing their research strategy, analysts are confronted with the interrelated issues of deciding “who is an entrepreneur” and identifying data sources suitable to analyze entrepreneurial behavior. In many instances, the focus has been on those individuals who identify themselves as self-employed in large cross-sectional data sets like the Census or Current Population Survey (CPS), or panel data sets such as the Panel Study of Income Dynamics (PSID), Survey of Income and Program Participation (SIPP), National Longitudinal Survey (NLS), or Health and Retirement Study (HRS).

While the existing empirical literature on the nature and causes of entrepreneurship is interesting and informative, it remains the case that woefully little is known about who really is an entrepreneur, and how sensitive earlier results are to variations in the definition of entrepreneurship. Survey data sources often provide a wealth of demographic and financial data for a large number of respondents, but bring with it uncertainty over the ability to define and describe the sample of entrepreneurs. It is valuable to learn if the results of earlier studies may be applicable to other definitions of “entrepreneur,” either more narrow or more broad.

To address this, we examine trends in various forms of entrepreneurship via a large sample of individual tax return data. While tax returns also carry some drawbacks, they provide a reliable source of microdata on entrepreneurship. The use of tax-return data also presents researchers with multiple indicators of business activity, each of which is a potential means of classifying entrepreneurs.

For purposes of building a “bridge” to related studies, our objective herein is to examine how the choice of definition for “entrepreneur” affects the estimated levels and dynamic patterns of entrepreneurship using tax data. In particular, we focus on the role of access to capital in influencing the dynamics of entrepreneurship. Additionally, a number of important tax policy changes during the 1980s might have affected entrepreneurial activity. Perhaps most important was a dramatic change in the relative payroll tax (i.e., FICA and SECA) treatment of the self-employed vis-à-vis employees (see Bruce [2000]). Further, changes in marginal income tax rates such as those in the Tax Reform Act of 1986 might have altered after-tax returns to entrepreneurship. While we do not formally control for the effects of taxes in the analysis that follows, we consider their importance in discussing our primary results.
The remainder of the paper is organized as follows. In Section I, we document the rising rate of entrepreneurship over the 1980s (1979 to 1990, to be precise) and examine the robustness of the estimated pattern with respect to the inclusion or exclusion of sole-proprietorship, partnership, sub-chapter S corporation, rental/royalty, and farm activity. We also examine the degree to which the rising trend in entrepreneurship can be “explained” by factors related to tastes for risk or access to capital. Section II focuses on the individual-level dynamic relationships among alternative indicators of entrepreneurship to see whether there exists a canonical “life-cycle” pattern to indicators of entrepreneurship. The final section is a summary with suggestions for future research.

To anticipate the bottom line, our examination of tax-return data indicates that the choice of definition for entrepreneurial activity has little impact on the overall conclusions. The various indicators that we examine tend to display similar trends over time. Access to capital or debt can have a positive effect on overall entrepreneurial activity, primarily through enhancing entry probabilities but also by reducing exit rates. Finally, a canonical progression of entrepreneurial states does not emerge from the data. In other words, entrepreneurs tend to focus on one particular form (e.g., self-employment, partnership, etc.) for this business, rather than moving through a variety of organizational forms.

I. ENTREPRENEURSHIP: TAX-BASED IDENTIFIERS
We begin our investigation by examining trends in entrepreneurship as revealed in tax data.

I.A. Data
Our data are drawn from the University of Michigan Tax Research Database. In constructing this panel data file, the Office of Tax Policy Research (OTPR) at the University of Michigan acquired the public-use tax return data released by the Internal Revenue Service (IRS) Statistics of Income (SOI) Division and converted them to user-friendly format. The 1979-1990 panel is constructed from annual IRS-SOI Individual Tax Model Files, which contain up to 200 pieces of information for between 80,000 and 250,000 personal income tax returns in each year. Within each Individual Model File is a subset of returns that were randomly selected to be part of a panel of taxpayers whose returns would be drawn year after year. In total, the panel includes data from over 200,000 tax returns. Approximately 6,000 filers are present in the panel for all 12 years.

I.B. Trends In Alternative Measures Of Entrepreneurship
As mentioned above, taxpayer data provide a number of categories of entrepreneurial activity. We focus on those with sole proprietorships (as evidenced by the presence of a Schedule C), partnerships, subchapter S corporations, and rental, royalty, or farm income. Survey data typically capture the first three of these as “self-employed,” but researchers often omit those in the latter categories as “partially” entrepreneurial. As a full analysis of each of these individual types would be rather unwieldy, we combine them into a smaller number of categories of entrepreneurial activity.

Specifically, we create four indicators for entrepreneurship, listed in increasing order of breadth. The first includes those who file a Schedule C, the second also includes those with partnership or subchapter S corporation income, the third adds in those with rental or royalty income, and fourth adds those filing Schedule F for farm income. It quickly became clear that the two narrowest measures performed very similarly over time, as did the two broadest measures. With this background, we focus our discussion on the second and third measures of entrepreneurship, which we refer to as our “narrow” measure and our “broad” measure, respectively.
The starting point for our discussion of the data is Figure 1, which shows the evolution over the 1980s of measures of entrepreneurship. As the figure makes transparent, the tax-based data display a general rise in entrepreneurship, with the trend most pronounced among joint returns (versus single returns).\(^3\) Looking more closely, one sees that for single returns the narrow definition rises overall, but the increase is far from uniform. Expanding to a broad definition shows first an increase, followed by a decline toward the end of the decade. For joint returns, the rise is fairly uniform for both measures.

An interesting result is that none of the trends exhibit increased volatility around key tax policy changes. Despite a leveling of the payroll tax playing field between wage and self-employment in 1984 and a marginal income tax rate reduction in 1986, the general increase in entrepreneurial activity is uninterrupted.\(^4\)

Year-to-year changes in entrepreneurship follow (by definition) from shifts in the rates of entry into entrepreneurship and exit from entrepreneurship. Indeed, the steady-state relationship between the rate of entrepreneurship, \(\pi\), the entry rate, \(e\), and the exit rate, \(x\), is given by

\[
\pi = \frac{e}{e + x}.
\]

Accordingly, one would expect that increases in the rate of entrepreneurship will be driven by either increases in the rate of entry, or decreases in the exit rate (increases in survivorship). Of course, equation (1) characterizes the long-run relationship. In the short-term, fluctuations in the components of (1) may not balance precisely.

To investigate this matter, Figures 2 and 3 show the corresponding yearly entry and exit rates based on our tax data.\(^5\) Beginning with Figure 2, one sees that over the course of the 1980s entry rates rose uniformly, regardless of the extensiveness of our definition or the filing status. The entry rates also seem to react to the increased relative payroll taxation of sole-proprietors enacted in 1984—all four series show a slight decrease in entry around this time. Nonetheless, the general upward trend is consistent with the overall increase in entrepreneurship shown in Figure 1. In comparison to Figure 1, however, entry rates are “noisier” and display a great deal of year-to-year fluctuation.

This characteristic, \(i.e.\) “noise”, is even more pronounced for exit rates (see Figure 3). However, Figure 3 has two features that are quite different than the previous two figures. First, exit rates for single returns are higher than for joint returns (the opposite is true for both entry rates and the overall rate of entrepreneurship). Second, in this instance the definition of entrepreneurship is more important. While exit rates fluctuate considerably, over the 1980s exit rates fell for three of the four possibilities. Note, however, that for the narrow definition of entrepreneurship among single returns, exit rates rose. Finally, exit rates do not display an upward spike around the key payroll tax reform of the mid-1980s. This lack of a response is perhaps partially responsible for the more stable increase in overall activity as shown in Figure 1.

The overall changes are made clearer when we consider the net changes over the sample period in entrepreneurship rates and the constituent dynamics for single and joint returns. Regardless of the definition of entrepreneurship, entry rates increased and exit rates decreased over time for joint returns, leading to an overall increase in entrepreneurship. This is perhaps a result of the general increase in self-employment among women, and married women in particular (see Devine [1994] or Bruce [1999]). Results for single filers are more sensitive to definitional issues, as the exit rate falls by a larger amount when the broader measure is used. Overall entrepreneurship rates and entry rates rise by similar amounts over time for single filers regardless of the definition. This general finding leads us to focus on our narrow definition of entrepreneurship for the remainder of the analysis.
I.C. Has There Really Been a Rise in Entrepreneurship?

The results presented above suggest a simple story: entrepreneurship was on the rise during the 1980s, driven both by increased entry of new entrepreneurs and enhanced survival of those already in business (although the evidence is a bit less clear-cut on the latter point). But why? Much of the popular discussion focuses on an increased “taste” or proclivity for entrepreneurship in the United States. In contrast, the bulk of the econometric research has focused on the degree to which a variety of “constraints”—taxes, access to capital, health insurance, discrimination—have determined rates of entrepreneurship over time, across space, or among races. To be sure, income tax data provide precious little information on basic demographic information and underlying wealth. We make use of suitable proxies where appropriate.

To investigate the underlying causes of increased entrepreneurship, we run reduced form regressions controlling for changes in the age-structure of our tax returns, the marital structure, and the number of child-based exemptions claimed. These regressions, in effect, control for variations in the willingness to bear risk associated with these demographic factors, and do so in a non-parametric way that imposes little structure on the data.

In the same way, we also control for access to capital via a mortgage or personal wealth using indicators for the mortgage interest deduction and the presence of capital income on the tax return. In each instance, we examine the contribution of each risk or capital factor to the overall trend in entrepreneurship.

Specifically, we estimate the following regression

\[ y_{it} = \alpha_i + \beta_A A_{it} + \delta_E E_{it} + \gamma_C C_{it} + \phi_D D_{it} + \epsilon_{it} \]

where \( y_{it} \) is an indicator variable for whether tax return \( i \) shows entrepreneurial activity in year \( t \), \( A_{it} \) is an indicator variable for whether the return shows an exemption for age greater than 65, \( E_{it} \) is an indicator variable for the presence of dependent exemptions (specifically, either children at home or away from home), \( C_{it} \) is an indicator variable for capital income, and where \( D_{it} \) is an indicator variable for mortgage interest. We estimate (2) separately for single and joint returns. In addition, we re-estimate (2) using as the dependent variable an indicator for entry into entrepreneurship and exit from entrepreneurship, respectively.

Notice that our specification of (2) permits the demographic variables (age and exemptions) to have differing impacts by year and by filing status. Thus, the estimates of the coefficients provide a flexible summary of their influence on the evolution of entrepreneurship in the 1980s. In a similar fashion, and the focus of more attention in what follows, we view the variables \( C \) and \( D \) as providing an index of entrepreneurs’ access to equity capital and debt. Thus, our estimated coefficients summarize the association of the changing nature of these financing constraints with entrepreneurship.

Finally, consider the role of the \( \alpha \) terms in equation (2). These capture the “pure”—that is, corrected for demographics and capital access—yearly shifts in rates of entrepreneurship, entry and exit. Thus, we can think of these estimates as “regression-adjusted” trends in entrepreneurship.

With this in mind, consider Figure 4, which displays for joint returns a comparison between the raw rates of entrepreneurship and our regression-adjusted estimates of the yearly shifts. Among joint returns, it is clear that the overall level of entrepreneurship is intimately tied to the demographic and capital structure—this is no surprise. Interestingly, however, the overall rise in the raw rate (7.1 percent) exceeds our estimate of the adjusted rate (5.5 percent). Put differently, the figure suggests that not all of the rise may be attributed to tastes or societal shifts toward entrepreneurship.

Is the same pattern present in the single returns? In Figure 5, one finds that the overall rise in entrepreneurship is 2.7 percent. Adjusting via our regression strategy shows that over the
1980s the residual increase is only 2.0 percent. Again, at least part of the rise in entrepreneurship must be attributed to changes in demographics or capital structure.

If so, one would expect to find the same impacts on entry and exit rates. Hence, we show in Figures 6 to 9 the comparison of raw versus adjusted entry and exit rates. What lessons are present in the Figures? Consider first the evidence on entry. One finds that adjusted entry rates show considerable fluctuation and rise more than raw data for both joint and single returns. This can be interpreted either as evidence of a slight increase in risk-taking in the data, or perhaps a response to more general trends in business structure. For example, there was a marked increase in outsourcing during the 1980s which led to increased numbers of independent contractors. This may appear as a rise in those classified as “entrepreneurial” in our tax data.

In contrast to the entry rates, while the adjusted exit rates show evidence of shifts over the 1980s, there is little change in the raw rates. And, to the extent that changes are present, the movements in raw rates are much smaller than in the adjusted data. This suggests that changes in the observable factors helped to diminish the volatility caused by unobservable factors.

The discussion thus far has focused on the “taste” for entrepreneurship and the “skill” for survival—some of the unobserved components of overall entrepreneurship that influence the rates of business activity, but are not shown in observable ways on the tax return. We turn now, instead, to looking at the direct influence of observables on the rates of entrepreneurship. In keeping with previous literature, we focus on the role of capital and debt access in the remaining Figures.

To get a flavor of the analysis, consider Figure 10. In the figure, we use our estimates of the impact of $C$, by year, to show the difference between the adjusted rate of entrepreneurship with $C_{it}=0$ and the adjusted rate with $C_{it}=1$. As the figure makes clear, the association between entrepreneurship and capital income is quite strong throughout the sample period, a result that is echoed among the single returns (see Figure 11).

An equally interesting result from Figures 10 and 11 is that entrepreneurship rates increase over time for both categories—with and without access to capital. Figures 12 and 13 take a closer look at entry rates by capital access, and reveal substantially more volatility as observed previously. Those with access to capital are consistently more likely to enter, but entry rates tend to trend upwards only for single filers over time. Somewhat surprisingly, the largest upward trend in entry is observed among single filers without access to capital.

As shown in Figure 14, exit rates trend upward over time for joint filers. The patterns are not as clear for single filers (see Figure 15). For each of these figures, though, access to capital translates into a much lower exit rate from entrepreneurial activity. Similar analysis with our broad definition of entrepreneurship yielded nearly identical patterns. In the end, access to capital has the expected effect of increasing entry rates and decreasing exit rates, regardless of filing status or entrepreneurial activity definitions.

We performed a similar analysis for the access to debt, as measured by the presence of mortgage interest deductions. This exercise revealed patterns of overall entrepreneurial activity and entry that are similar to those for the analysis of access to capital and do not warrant lengthy additional discussion. In sum, access to debt has the expected effect of increasing entry and overall entrepreneurship rates.

The effects of access to debt on exit rates were far less intuitive. Access to debt seems to have little effect on exit rates when the narrow definition is used, but a slight negative effect on exit rates when the broad definition is used. These patterns are similar but far less uniform for single filers. While access to capital can enhance both entry and survival, debt appears only to help in the start-up phase of entrepreneurial activity.
II. THE ENTREPRENEURIAL LIFE-CYCLE: A TAX PERSPECTIVE

Thus far we have noted that alternative, tax-based measures of entrepreneurship give similar (if not identical) insights into the overall prevalence and year-to-year variation in entrepreneurship. Moreover, the movements into and out of entrepreneurship in these tax-based data appear to have the same relationships to capital-access and the ability to borrow as found in other settings. These findings provide support for the notion that tax-based data accurately reflect entrepreneurial dynamics in the economy.

Tax data have advantages over other data sets, however. In particular, our tax-return data show a variety of forms for business activity: sole-proprietorships, partnerships, sub-chapter S corporations, rental businesses, and so forth. Hence, they allow us to examine a distinct, but closely related, question: does there exist an individual-specific sequence of indicators of entrepreneurship? That is, is there a life-cycle to the kinds of business forms that an individual will adopt?

To address this question, we use our panel data to compute empirical estimates of the individual-specific probability of making transitions from one business form to another. To be concrete, consider Table 1, which shows the 8x8 transition matrix. To construct the table, we classified returns on the basis of the business activity: (1) none, (2) sole-proprietorship only, (3) sole-proprietorship and partnership or S-corporation, (4) sole-proprietorship and partnership or S-corporation and rental or royalty (“all”), (5) sole-proprietorship and rental or royalty, (6) partnership or S-corporation only, (7) partnership or S-corporation and rental or royalty, and (8) rental or royalty only. To the extent that there is a dominant dynamic sequence of business forms, one would expect the individual-level transitions to follow a common path. If, on the other hand, choices for business formats are driven by idiosyncratic matters, one would expect an equal distribution in the table.

What do we learn from Table 1? Each entry shows the fraction of returns in the corresponding row that make a transition to the status shown in the corresponding column. Thus, for example, among single returns 1.1 percent of individuals who had no business activity made a transition to sole-proprietorship the next year, while the corresponding rate was 2.7 percent among joint returns. Or, to take another example, 0.96 percent of single sole-proprietors took on partnership or S-corporation activity in the next year; the corresponding rate for joint returns was 1.1 percent.

The patterns in Table 1 suggest several themes. First, the diagonal entries show that the dominant transition is “more of the same.” Thus, despite the notoriously high exit rates of small firms, the data contain considerable year-to-year correlation in activities.

Second, the next most likely transition is to “exit entirely.” That is, even when there are multiple business activities—in the extreme “all”—the second most likely transition is to have “none” in the next year—roughly 20 percent among both single and joint returns. This is quite a striking result.

In similar fashion, the third observation is that “purists”—those engaged only in a sole-proprietorship, partnership, S-corporation, or rental business—are “all or nothing.” That is, the sum of the probabilities of staying the same and exiting to “none” is higher than for any other possibilities.

Lastly, to the extent that a there is a transition, it usually takes the form of adding a new business activity. Thus, we don’t see the transformation of a single business to a new form, but rather the addition of other business forms. This latter observation, in particular, militates against a “growth chain” view in which a start-up business moves deterministically through a series of formats.

Our final exercise is to merge the insights regarding the importance of capital and debt access with our emerging picture of transitional dynamics. Specifically, in Tables 2 and 3 (for
single and joint returns, respectively) we compare the transition rates for those without access to capital or debt (top entry in each row) to those with access (bottom entry in each row).

We leave a detailed inspection of each table to the reader. For our purposes it is interesting to note that the overall pattern of transitional dynamics is similar for both groups, and thus reflective of the patterns detailed in Table 1. However, previous research has shown a strong link between access to capital and the rates of entry and survival. These tax-based data support this view, as the entry rates are higher for the bottom entries nearly uniformly, while exit rates are lower.

III. SUMMARY AND CONCLUSION

Our analysis of a 12-year panel of individual taxpayer data has revealed a number of important results. First and most importantly, the choice of definition of “entrepreneur” has little if any impact on the observed trends in entrepreneurship as well as the analysis of some of the underlying causes of those trends. In other words, researchers may not need to worry as much about the ability of self-employment or other indicators in survey data to capture entrepreneurial activity.

A second key result is that access to capital and debt have the expected effect of increasing entry rates and reducing exit rates. Nonetheless, regression results indicate that demographics and a general increase in risk-taking are also responsible for part of the general rise in entrepreneurial activity during the 1980s.

Finally, no clear life-cycle pattern of entrepreneurial activity emerges from a multi-state transition matrix analysis. Those with one form or another of entrepreneurship tend either to stay in that form or to exit entirely. To the extent that movements among forms take place, they typically involve the addition of another form rather than a reduction. Again, though, these types of transitions are clearly the minority of all transitions.

A number of important questions remain for future research. For example, the relative importance of tax policy must be explored in a more exhaustive multivariate framework. Further, additional descriptive detail regarding entrepreneurial finance can be gleaned from these data, especially for those filing a Schedule C.
REFERENCES


FOOTNOTES

1. The trend is not restricted to the United States. Andrews [1998, p. D1] reports “In a historic switch that has both amazed and alarmed financial analysts, investors from Frankfurt to Brussels to Amsterdam have fallen in love with entrepreneurs… Compared with the attraction in the United States…the infatuation here is in its early stages. But compared with Europe’s past, it is a revolution.”

2. See, for example, Dunn and Holtz-Eakin [2000] for the NLS, Bruce [2001 and 2000] and Gentry and Hubbard [2000] for the PSID, Schuetze [2000] for the CPS, or Bruce, Holtz-Eakin, and Quinn [2000] for the HRS.

3. We compress five filing status options into two categories—single and married—as follows: unmarried heads of household and widow(er)s with dependent children are combined with single filers, while married couples filing separately are combined with those filing jointly.

4. Of course, this is not a substitute for a structural analysis of the impact of taxes on entrepreneurship. See Bruce and Holtz-Eakin [2001].

5. We define entry as having the particular form of entrepreneurial activity on this year’s tax return, but not last year’s. Exit is defined in a similar fashion.

6. See Bruce [2001 and 2000], Bruce and Holtz-Eakin [2001], Carroll, Holtz-Eakin, Rider, and Rosen [Forthcoming, 2000a, and 2000b] for recent studies of the role of taxes; Holtz-Eakin, Joulfaian, and Rosen [1994a and 1994b], Bruce, Holtz-Eakin and Quinn [2000] on the role of capital market constraints; and Holtz-Eakin, Penrod, and Rosen [1996] or Bruce, Holtz-Eakin, and Quinn [2000] on the effects of health insurance. This self-serving list is not meant to be exhaustive; more extensive references are contained in each of the articles.

7. Regression results underlying Figure 4 (and subsequent figures) are available from the authors upon request.