Entrepreneurial Spawning of Scientists and Engineers: Stars, Slugs, and Small Firms

Daniel W. Elfenbein, Barton H. Hamilton, Todd R. Zenger*

December 10, 2007

Abstract

We examine the determinants of transitions from paid employment into self employment by scientists and engineers between 1995 and 2001. We find that those working in small firms are significantly more likely to become self employed than those working in large firms. Entrepreneurs coming from both small firms and from large firms are more likely to be high performers (stars) or low performers (slugs) as measured by their pay in their prior jobs; this finding is particularly pronounced among those leaving small firms. Finally, we find some evidence that entrepreneurs coming from small firms perform better than those who come from large firms: they are more likely to persist in self-employment and earn more in their first period of self-employment, controlling for a number of other factors. We explore the degree to which these relationships can be explained by theories that focus on differences in opportunity costs, the strength of pay - performance relationships, and diversity of activity sets, and the acquisition of entrepreneurial capital in firms of different size.

PRELIMINARY DRAFT—PLEASE DO NOT CITE
WITHOUT AUTHOR'S PERMISSION

*Olin Business School, Washington University in St. Louis

We would like to acknowledge the National Science Foundation for access to the data. The use of NSF data does not imply NSF endorsement of the methods or conclusions described in this report.

1. Introduction

Entrepreneurship has been lauded by numerous observers as a driving force behind economic growth and technological change, and not surprisingly, the topic has attracted a great deal of theoretical and empirical attention. A large body of research has focused on the determinants of entrepreneurship at the level of the individual, examining the role of such factors as gender, race, education, and credit constraints on individual's decisions to start their own firms (e.g., Evans and Leighton, 1989, Borjas and Bronars, 1989, Evans and Jovanovic, 1989; Blanchflower and Oswald, 1998, Hurst and Lusardi, 2004), as well as the importance of human capital and non-pecuniary returns (e.g., Lazear, 2005; Hamilton, 2000). In contrast to the large numbers of studies in these areas, substantially less attention has been paid to the role that employers play in "spawning" entrepreneurial ventures. As a large number of entrepreneurs move into these pursuits from paid employment, this represents a significant gap in the literature. Of particular interest is how employer characteristics affect entrepreneurial activity and the performance of resulting entrepreneurial ventures. Gompers, Lerner and Scharfstein (2005), hereafter GLS, begin to fill this gap, examining the relationship between public corporation characteristics and the formation of venture-backed startups. In this paper we extend the investigation of entrepreneurial spawning to a broader set of employers, both publicly traded corporations and private employers, as well as to a broader range of startups, both large and small, founded by scientists and engineers.

Scientists and engineers are a particularly important group to study given their importance to the development of new, innovative technologies. Using panel data on scientists and engineers covering 1995-2001 obtained from the National Science Foundation's Scientist and Engineers Statistical Data System (SESTAT), Table 1 highlights the relationship between employer size and self-employment entry that has received surprisingly little attention in the entrepreneurship literature[1]: For both PhD scientists and engineers, and those with Bachelor/Master degrees ("non-PhD"), very small employers appear to

[1] To our knowledge, reports of this relationship are limited to Boden (1996) and Wagner (2004).

spawn a vastly disproportionate number of entrepreneurial startups. For example, the last column of the table shows that among PhDs working in a small (1-25 employees) firm in 1995-1999, 18.7% move to self-employment within the next two years, compared to only 0.9% of individuals at firms with 5000 or more employees. While turnover is substantially higher at small firms, roughly one in two small firm leavers enter self-employment, compared to fewer than one in 10 who leave a very large firm. GLS note a similar though less dramatic relationship between firm size and spawning, but their data set consists of public corporations and conditions on successful (i.e., venture backed) entrepreneurs. Table 1 suggests that this phenomenon is observed more generally.

The goal of this paper is to determine the extent to which the empirical phenomena described in Table 1 can be explained by a variety of models of the labor market and entrepreneurship. These explanations fall into three categories. First, models of occupational choice and sorting (e.g., Roy, 1951) emphasize heterogeneity in workers' sector specific abilities. It has been argued that small firms attract lower ability workers (Troske, 1999), suggesting that small firms may spawn more entrepreneurs because their employees have lower opportunity costs in paid employment.[2] Conversely, both GLS and Hellman (2007) suggest that the best employees at large firms may become entrepreneurs because these bureaucratic firms are less able to tie pay to performance. Star employees then turn to self-employment to fully capture returns to their ability. Second, human capital models suggest that individuals working in small firms accumulate skills that are valuable in entrepreneurship. For example, a worker in a small firm may be forced to be a "jack of all trades," which Lazear (2005) argues is a key characteristic of successful entrepreneurs. In a related vein, GLS argue that individuals employed in "entrepreneurial" firms are exposed to networks of suppliers of labor and capital, as well as potential customers, which increase "entrepreneurial capital." Third, workers in small firms may have preferences for non-pecuniary job attributes, such as autonomy, that are even more pronounced in self-employment (e.g., the opportunity to "be your own boss").

[2] An alternative sorting story may be built on heterogeneous risk aversion among workers (Parker 2006).

We explore these theoretical explanations of the relationship between firm size and entrepreneurial spawning using a new dataset (the SESTAT) of science and engineering graduates from American universities between 1947 and 2001 developed by the NSF that contains extensive information on individuals' education, job experience, and demographic characteristics. More importantly, the SESTAT is especially suited for our analysis because it has longitudinal information from 1993 – 2001 for a large number of individuals. The large sample size is necessary to provide sufficient numbers of observations when analyzing relatively rare transitions in some cases (e.g., from large firms to self-employment).

The data that we examine are distinct from those that have been used in other studies of entrepreneurship and self-employment. Prior studies focus either on broad national samples, such as the National Longitudinal Survey of Youth, which attempt to represent the entire working population (e.g., Evans & Leighton) or on very narrow data sets of venture-backed start-ups (e.g., GLS 2005). Studies of the former type may overstate, from a policy point of view, the importance of small proprietorships since entrepreneurship is identified by activities ranging from barbershops, caterers, and convenience stores founded by those with limited education to venture-backed technology start-ups by PhD Studies of the later type, while highly valuable, draw conclusions based on examination of an elite group whose members may not be responsive to the same considerations as those of the broader population. The data in this paper occupy a middle ground between these types of studies. The individuals we examine have all achieved at least a bachelor's degree in a science and engineering field, and in several of our analyses, we focus only on those who have received PhD's in these fields. Our sample embodies those who are most likely to be the targets of policy-makers concerned with entrepreneurship as a force of economic growth—individuals with high levels of human capital in dynamic, knowledge-intensive fields. We consider this an important strength of our data.

Our preliminary empirical analysis yields a number of interesting results. First, we find that some, but not all, of the increased propensity of small firm employees to become self-employed can be explained by a different (and broader) set of activities that small-firm employees pursue on the job,

consistent with "jack-of-all trades" theories. Second, we find little support for the idea that differences in pay levels between small and large firms account for much of the "small firm effect." In the PhD sample we construct, the pay distribution varies little by firm size, yet the rate at which PhDs transition to self-employment from small firms is substantially higher than the rate from large firms. Third, we find that the future self-employed are significantly more likely to come from the top and bottom ends of the paid employment distribution, consistent both with prior findings that misfits or "slugs" enter self employment, and with selection-based models that suggest that the most able workers become self employed to maximize the returns to their human capital. Fourth, we find that entrepreneurs coming from small firms persist in self employment and "better" small-firm employees earn more than those from large firms, and that a disproportionate share of high-ability small firm employees start their own firms. This evidence is difficult to explain with ability-sorting logic alone. Rather we interpret it as suggesting that working in small firms provides the set of more able employees with "entrepreneurial capital" that may improve their ability to capitalize on opportunities. Consequently, both the likelihood that they will become self-employed and their success conditional entering self-employment increase.

We proceed with the paper as follows. Section 2 explores the related theoretical and empirical literature. Section 3 describes the data. Section 4 examines entrepreneurial entry and sorting by ability. Section 5 explores performance differences between entrepreneurs coming from small and large firms, and Section 6 concludes.

2. Literature

In exploring the role of ability and prior employment on entrepreneurship decisions, we first examine economic models that build on the matching logic of Roy (1951) and Jovanovic (1979), where individuals with differing levels of sector-specific abilities choose the employment or entrepreneurship state that yields the highest level of utility. A number of theories argue that high ability worker will sort into large firms, for reasons such as complementarities between ability and capital (Lucas, 1978); complementarities between highly skilled managers and highly skilled workers (Oi, 1983); or the fixed

costs associated with hiring high ability workers that are more easily borne by large firms (Kremer, 1993), Kremer and Maskin 1995). Consistent with this logic, empirical studies show that large firms pay more (Troske, 1999; Brown and Medoff, 1989), suggesting that workers in small firms have a lower opportunity cost of self-employment entry, all else equal.

Other scholars have emphasized that large, bureaucratic firms are less able to directly link pay to performance, perhaps because of higher measurement costs (Garen, 1985). Large firms therefore link pay to observable measures such as schooling, while small firms, by more effectively rewarding individual performance, attract the more able employees from large firms (Zenger, 1994). Thus, in the context of self-employment, the most talented workers at large firms may choose entrepreneurship because they are unable to fully capture the returns to their own ability at their employer. In a study of academics in science, Zucker, Darby, and Torero (2002) show that "star" scientists—those who publish early and attract the most citations, are those most likely to leave academia to join or found a firm. The mechanism driving departure is the inability of the university to compensate for increases in a scientist's quality measures. In addition, GLS (2005) suggest that managers of large established firms may be less able to evaluate innovative ideas because it falls outside the company's core competence. Hellman (2007) also shows that when employees have the choice of working on core tasks at the firm or exploring new ideas, the firm may refuse to develop the innovation in-house and force innovative employees into entrepreneurship. To the extent that these innovators are the most able employees at large firms, these models suggest that self-employment entrants will be drawn from the top of the pay distribution. In the case of small firms, one might expect more flexible pay arrangements allow the firm to retain innovative employees. However, an implication of Hellman (2007) is that large firms may have the capability to develop employee innovation as internal, "intrapreneurial" ventures. Small firms may not have these resources, forcing talented employees at these firms to become entrepreneurs.

A second category explaining transitions to entrepreneurship highlights the role that small firms may play in enabling employees to acquire human capital that is valuable in self-employment. For example, Lazear (2005) argues that entrepreneurship requires a diverse set of skills including both

application knowledge and a wide range of management skills. Entrepreneurs not only require an entrepreneurial idea, but they require a more balanced, jack-of-all-trades set of skills. Arguably, employment in a small firm requires the employee to acquire a range of skills that will be valuable in subsequent entrepreneurial ventures. Similarly, Gompers, Lerner, Scharfstein (2005) suggest that those employed in small entrepreneurial firms gain access to valuable networks critical to entrepreneurship. Within these firms, future entrepreneurs learn essential steps in founding a firm. Finally, in a study of academic entrepreneurship, Stuart and Ding (2006) find that movement into entrepreneurship is more likely when colleagues and co-authors have prior experience in entrepreneurship. This prior experience of colleagues provides knowledge and networks critical to commercializing knowledge. Each of these models suggests that context (in our case small firm employment) is an important determinant of whether workers acquire human capital that will increase their chance of success in entrepreneurship. Consequently, small firms may spawn more successful entrepreneurs than large employers.

A final set of explanations emphasizes the non-pecuniary returns associated with self-employment. Individuals may choose to become entrepreneurs because they value autonomy and want to be their own boss (Hamilton, 2000). Small firms are likely to offer a higher level of autonomy than large firms, suggesting that some individuals may be attracted to these firms prior to entering self-employment. One implication of this model is that self-employment entrants from small firms may be willing to accept lower returns after becoming entrepreneurs than those spawned by a large firm.

A small number of related papers deal with similar phenomena. Wagner (2004) finds evidence of a strong relationship between prior experience working in a small or young firm and the likelihood of starting a business in a random sample of German employees roughly a tenth the size of our own. Parker (2006) explicitly addresses the role of employer size in predicting transitions to entrepreneurship using a theoretical model that investigates the role that heterogeneity in risk tolerance may play. In this model workers have identical ability but differ according to their level of risk aversion, and large firms are more capable of bearing risk than small firms, so they offer more wage insurance to workers. As a result, less risk-averse workers choose the contracts offered to them by small firms. Exogenous shocks open new

opportunities for workers to start their own firm, and those who take these opportunities are more likely to come from the set of employees who chose to work in small firms in the first place. Anderson and Wadensjö (2006) use data from Sweden to examine the relationship between earnings in paid employment compared to earnings in self-employment. They find, as we do, that entrepreneurs are drawn disproportionately from the high and low ends of the wage distribution, but they do not explore the importance of employer size in generating this relationship. Similarly, Velamuri and Venkataraman (2005) find a u-shaped relationship between income in paid work and transitions to self-employment using the Panel Study of Income Dynamics.

In work most closely related to our own, Braguinsky and Ohyama (2007) examine job selection among scientists and engineers using the SESTAT data. They develop a model of job-matching in which workers learn about their ability over time. An attractive feature of their model is that it predicts that entrepreneurs coming from the upper part of the paid wage distribution will differ in the types of firms they found from those who enter entrepreneurship from the bottom end of the paid wage distribution. Moreover they report an interesting finding: those who enter entrepreneurship in businesses related to their education earn more when choosing self-employment. Finally, Sauermann and Cohen (2007) use the SESTAT data to examine the relationship between firm types, individual's reported desire for pay vs. job security, effort and innovation performance. They find that startups may have better innovative performance because they attract more productive inventors (rather than by offering a better innovative "environment").

3. Data

Constructing the Samples

We construct two samples of individuals with science and engineering degrees using data from the Scientists and Engineers Statistical Data System (SESTAT). This data file is comprised of responses

to three separate surveys—the National Survey of Recent College Graduates (NSRCG), the National Survey of College Graduates (NSCG), and the Survey of Doctoral Recipients (SDR). All surveys are restricted to respondents who earned a science or engineering degree (S&E) for the relevant survey. The sampling methodologies vary widely across each of these three surveys. For example, the sample population for the 1993 NSCG was chosen by the Bureau of the Census to be representative of all college graduates in all fields as of 1990; follow-up studies from this sample, which we use in this paper, conducted in 1995, 1997, and 1999 were restricted to individuals with a S&E degree or employment in a S&E field. The NSRCG sampled S&E degree recipients from the prior two-year window, and the SDR defined as its sample population all people who had received an S&E doctorate from a U.S. institution by the year preceding the survey. In these surveys as well individuals respond to multiple episodes.

One sample is restricted to those with the PhD degree and draws observations from the restricted files of the Survey of Doctoral Recipients (SDR). This data set includes observations for the years 1993, 1995, 1997, 1999, and 2001.[3] Although 2003 data are available for the SDR, we do not include these data in the present analysis. A change in the survey design in 2003 appears to generate a sharp spike in the number of respondents listing themselves as incorporated self-employed; to eliminate this potential source of error, we discard the 2003 data.[4] Also, since questions about employer size were not asked in 1993, our analysis mainly excludes observations from 1993. The core years of the PhD sample, then, are 1995, 1997, 1999, and 2001. In many but not all instances, individuals respond to multiple episodes of the survey allowing us to analyze changes in their employment over time. We make the following additional restrictions to eliminate sources of undesirable heterogeneity:

- To avoid problems of retirement, full-time education, and other choices about whether to enter or remain in the labor force, we eliminate all of those who are not in the labor force in each year

[3] For details, see http://sestat.nsf.gov/.

[4] In future work, we can establish which of the respondents listed this choice mistakenly, based upon their responses to other survey questions.

between 1995 and 2001 and further eliminate all of those under age 22 or above age 65 in any year between 1995 and 2001.

- To avoid some of the impact of family choices on self-employment, we focus only on males (who comprise 71.3% of survey responses).
- Since we use measures derived from annual salary in our analysis below, we wish to avoid confounding total pay with choices about working part-time vs. full time. Therefore, we eliminate from the sample all those who report working fewer than 30 hours per week and all those who report working fewer than 30 weeks per year.
- Because we want to focus exclusively on scientists and engineers, we eliminate all those whose PhD degree was not in a science and engineering field and further, we drop from our analysis any individual who also holds a professional degree (such as an M.D., J.D., DVM., etc.).[5]
- To avoid confounds due to currency differences, all respondents working outside the United States are excluded from the sample.

The non-PhD sample is drawn from the restricted files of the Scientists and Engineers Statistical Data System (SESTAT) for the years 1993, 1995, 1997, and 1999. As with the SDR data, since 1993 does not include data on firm size, this year was not included in the sample. As with the PhD sample, we make a number of additional restrictions to arrive at our final sample:

- All respondents who enter via the SDR are eliminated, as are all other respondents who have earned a PhD or a professional degree. This leaves only respondents whose highest degree is a bachelors' or masters' degree.
- Again, we focus only on males (who comprise 67.5% of survey responses). We eliminate from the sample all those who report working fewer than 30 hours per week and all those who report working fewer than 30 weeks per year, and eliminate those working outside the United States.

[5] Masters in Business Administration (MBA) degrees are not considered by the NSF to be professional degrees. So PhDs who hold MBAs are included in the sample.

- We eliminate all those whose highest degree was not in a science and engineering field.[6]

We use all survey responses meeting the criteria described above to generate Table 1. As discussed in the Introduction, Table 1 illustrates the likelihood that an individual working for an employer of a given type in year *t-2* has either changed jobs, labeled "turnover" in the table, or has become self-employed by year *t*. The turnover category is inclusive of the transitions to self-employment, and may include some instances of individuals who have become owners in the firms for which they now work. The table illustrates that the frequency with which government, education, and university / research employees transitioned into self employment is considerably lower than the frequency with which workers in for-profit firms entered self employment. Across the PhD and non-PhD samples on average 3.6% of workers in for-profit firms transitioned into self-employment in any two-year period. By comparison, government, secondary education, and university / research institute employees' transition rates were 0.7%, 1.1%, 0.8%, respectively. Although the decisions by workers in these types of organizations to enter self employment are interesting, workers who have chosen employment in them may be responsive to a somewhat different set of considerations in choosing self employment than workers in for-profit firms. Additionally, the low rate at which transitions occur provides limited data from which to identify the factors that impact these transitions. For these two reasons, we focus solely on workers in for-profit enterprise in the remaining analysis.[7]

Sample Characteristics

Tables 2a and 2b provide summary statistics on the PhD and non-PhD samples, respectively. The data are constructed from survey responses by the surveying agency. An analysis of the summary statistics reveals a number of differences across the two sample groups. Respondents in the PhD sample

[6] Science and engineering includes social science disciplines such as economics, sociology, and political science, but not finance, accounting, management, or humanities.

[7] In the PhD sample, workers in for-profit enterprise account for roughly three in ten of the responses of males in the survey. In the non-PhD sample, workers in for-profit enterprise account for roughly two-thirds of males' responses.

have a median age of 44 and a median annual salary of 83,000 compared to values of 36 and 52,000, respectively in the non-PhD sample.[8] Those in the PhD sample have a median tenure of 4 years and one month in their current job, compared to a median tenure of 2 years 10 months in the non-PhD sample. PhD respondents report a higher level of median hours worked, 48, than those in the non-PhD sample, who report a median work week of 45 hours. Respondents in the PhD sample are more likely to have their highest degree in the life sciences and are less likely to have their highest degree in computer science or engineering than respondents in the non-PhD sample. Individuals in the PhD sample are more likely to be married, are more likely to have a non-working spouse, and have more children on average living in the household than those in the non-PhD sample. The self-employment rate of the PhD sample is higher (12.7% vs. 8.9%) as is the likelihood of working in firms of 5000+ employees (conditional on being employed in for-profit business), and the rate at which those in the PhD sample change employers is higher (22.3% vs. 17.7%). PhD sample respondents are more likely to list R&D as their primary activity in the workplace, and are more likely to declare that various types of R&D activity, consumed 10% or more of their working hours. The non-PhD sample, by contrast was more likely to be engaged in computer applications; accounting, purchasing and contracts; production, operation and maintenance; quality or productivity management; and sales, purchasing, and marketing. Both samples reported similar levels of engagement in employee / human relations, and in managing or supervising people. The geographic distribution of the two samples was highly similar with two exceptions. Respondents in the non-PhD sample were more likely to be located in mid-Atlantic states, and respondents in the PhD sample were more likely to be located in Pacific states.

[8] Two salary figures are reported. Salary figures represent annualized salaries before deductions and were constructed by surveying agencies by combining information on weekly / monthly earnings and weeks / months worked. The first salary figure reported comes from the confidential data file and represents, supposedly, the full distribution of salaries among respondents. We are skeptical that this measure reports large salary figures accurately for respondents to the NSRCG and NSCG surveys given the large mass of observations at the $150,000 level; a similar mass of observations on $150,000 is not a characteristic of the SDR responses. The second salary figure comes from the unrestricted data file. This annualized salary measure has been intentionally top-coded by the surveying agencies at $150,000. In regressions employing the PhD sample only, we employ the first measure, but due to our concern about selective censoring in the data, we use the top-coded measure in analyses of the non-PhD data and in all analyses that pool the two samples.

Table 3 compares the means of several of the key explanatory variables across different employer types. In this table, we include all self-employed, not just those who transition into self-employment while under observation. A notable feature of the PhD sample is the limited difference in average salary across firm size categories. On average, PhDs in firms of 1-25 employees earn slightly more than those in the firms of size 5000 +; the difference in the top-coded figures, however, indicate that the median earnings in the smallest firms are lower than in the largest firms (see footnote X above for a discussion of top-coding in the sample). Among non-PhDs, where salaries are top-coded, the comparison suggests that workers do earn more in larger firms. The comparison of activities by firm size is also notable. Workers in small firms are likely to be engaged in a broader range of commercial activities and a smaller range of R&D activities than workers in small firms. Those who are self-employed are engaged in the broader range of commercial activities, on average, than those working in firms of any size and seem less likely to pursue R&D activities.

4. Entrepreneurial Entry and Sorting

Transitions into Self Employment

In the samples we construct above, nearly half of all movement into self-employment comes from firms of less than 25 employees and just under two-thirds comes from firms with fewer than 100 employees. Given that such a disproportionate share of all movement into self-employment comes from small firms, a critical empirical question is explaining this simple fact. We begin by examining the factors that are correlated with individuals' transitions from paid employment to self-employment. In particular, we are interested in understanding the degree to which the strong relationship we observe between firm size at time *t* and the likelihood of being self-employed at time *t+2* can be explained by other independent and contextual factors. To explore this question, we estimate the following equation:

$$PR(y_{i,t+2} = 1 \mid y_{it} = 0) = \alpha + \beta \mathbf{X}_i + \gamma \mathbf{Z}_{it} + \mu_t + \varepsilon_{it} \qquad (1)$$

In this equation, y_{it} equals 1 if individual i is self employed in year t and 0 otherwise. The vector $\mathbf{X}_i$ is a set of time-invariant individual characteristics such as race, country of origin, and field of highest degree[9], and the vector $\mathbf{Z}_{it}$ is a vector of all potentially time-varying individual characteristics, such as marital status, number of children in the household, and location, as well as all characteristics of the individual's employer, job activities, and the individual's job tenure and pay at his particular employer at time t. Employer characteristics within $\mathbf{Z}_{it}$ include firm size and location (generally region), job activity variables include dummy variables indicating whether the employee's principle job is research and development, a measure of the diversity of activities pursued on the job, as well as fourteen dummy variables that reflect the activities on which the individual reported spending 10% or more of his time in any given week. Differences in the average rate of transitioning into self-employment over time are captured by μ_t and ε_{it} represents the idiosyncratic error. We estimate equation (1) only for those who are paid employees at time t; i.e., self-employed individuals are excluded from the estimation. The estimated coefficients can be interpreted as the likelihood of transitioning into self employment at $t+2$ as functions of $\mathbf{X}_i$ and $\mathbf{Z}_{it}$, rather than the likelihood of being self employed given $\mathbf{X}_i$ and $\mathbf{Z}_{it}$.

In Table 4a, we report estimates of equation (1) using a probit specification for the PhD sample. To facilitate interpretation, we display marginal effects coefficients. Column 1 serves as a baseline for considering the importance of employer size at time t in explaining self-employment at $t+2$, controlling only for year effects. The results reflect the patterns evident in Table 1. Employees of smaller firms transition into self-employment much more frequently than those working in larger firms. The likelihood of transition declines monotonically with our firm size categories. The differences in transition rates across firms are economically significant, with individuals in firms of size 25 employees and below transitioning into self-employment at a rate that is nearly six times the average rate in the sample. Individual and joint tests of equality across the firm size coefficients reject at the $p < .001$ level.

[9] In principle, the highest degree and the field of the highest degree could vary across years in the non-PhD sample, but since we exclude those who are not working full time from the sample, in practice there is no case in which these variable change during the course of our panel.

Column 2 adds a number of individual characteristics, the employer's location, and the individual's tenure in the current job. The estimated marginal effects on the firm size dummy variables decline in magnitude by 25 to 33%, but they remain economically and statistically significant. We observe significant differences in the rates at which PhDs in different fields transition to self-employment, in particular those with degrees in life sciences and social sciences are more likely to transition into self-employment. These differences remain as we add more covariates. Surprisingly, quadratic functions of age and job tenure are not significant in this regression or in subsequent specifications that employ a richer set of covariates.

The higher rate of entrepreneurial spawning by small firms is perhaps explained by the broader array of tasks that employees perform and therefore the broader knowledge accumulated by employees in small firms. To test this possibility, we introduce a set of 14 activity dummy variables in column 3. Although the activities of the employee on the job have important explanatory power, firm size effects remain. Introducing these covariates improves the fit of the model and further reduces the size of the coefficients on firm size by 8 to 22%, for a total reduction in magnitude of 31 to 42%. In unreported regressions, we explore the jack-of-all-trades theory of Lazear (2004, 2005) using different parameterizations of these 14 dummy variables. We find those whose jobs involve more commercial activities are significantly more likely to become self-employed but that those who pursue a wider range of research activities are less likely to become self-employed.

Recall from the discussion in Section 2 the argument that small firms will spawn more entrepreneurs because they pay less and so the opportunity cost of self-employment entry is lower. To assess this explanation, in Column 4 we depart from equation (1) slightly, modifying it to include the predicted paid-employment wage at t+2, while also correcting for selection effects influencing wages in paid employment. We estimate

$$PR(y_{i,t+2} = 1 \mid y_{it} = 0) = \alpha + \beta \mathbf{X}_i + \gamma \mathbf{Z}_{it} + \delta_1 E[W^{pe}_{i,t+2}] + \delta_2 E[W^{se}_{i,t+2}]) + \mu_t + \varepsilon_{it}, \qquad (1a)$$

where $E[W^{pe}_{i,t+2}]$ is the Heckman-corrected predicted wage for individual *i* at *t+2* should that individual have remained in paid employment and $E[W^{se}_{i,t+2}]$ is the Heckman-corrected predicted wage should the

individual enter self-employment. Following an approach similar to Willis and Rosen (1979), the selection equations are estimated as in equation (1) above (in the case of predicting the paid-employment wage we modify the equation to select on *remaining* in paid employment), and the wage in the second stage is estimated using a set of covariates in $\boldsymbol{X}_i$ and $\boldsymbol{Z}_{it}$ including employer size and wage at time *t* but excluding marital status and number of children, which are likely to affect the self-employment decision but not paid employment wages directly. We interpret the predicted paid employment and self-employment wages as two factors that determine the opportunity cost and expected return associated with entering self-employment. Opportunity cost considerations predict that δ_1 will be negative, while higher expected returns in self-employment should increase entry. In regressions using data from the PhD sample, the signs of the coefficients are as expected but are not significantly different from zero. It may be the case that predicted pay in t+2 is a poor predictor of the future stream of returns in each sector. Moreover, we note that the inclusion of the predicted wage does not change the small firm effect, implying that the relationship cannot simply be explained by differences in opportunity cost.[10]

Columns 5 through 7 begin to explore ability sorting explanations by examining the actual relationship between compensation in paid employment and the transition to self employment. To avoid estimating a supply response to wage rates, we employ the respondent's weekly wage (reported annual salary divided by reported weeks worked) as the focal measure of pay. We interpret the level of the weekly wage at time *t* after controlling for obvious correlates such as age and job tenure as providing a measure of the ability of the worker in paid employment. Although many factors determine pay, we assert that on average in this focused sample, more productive workers are likely to be paid more than less productive ones. In column 4, we enter the log of the weekly wage at time *t* directly into the estimation and find no significant correlation with the probability of transition to self-employment;

[10] In an unreported regression, we examine the log of the difference between the predicted paid employment and self-employment wage. The estimated coefficient on this difference has the predicted sign but is not significant at conventional levels.

moreover, incorporating the log weekly wage in the regression equation has little impact on the firm size coefficients.

In Columns 6 and 7 we explore the possibility of a non-linear relationship between one's prior position in the pay distribution and transition to self-employment. Specifically, we examine an individual's categorical position within the distribution of pay and the likelihood of transition to self-employment. Column 6 incorporates the individual's pay quintile into the transition equation. Individuals in the lowest pay quintile are significantly more likely to transition into self-employment as are individuals in the highest pay quintile. These differences are economically meaningful as well. Relative to the middle quintile of pay, being in the lowest quintile raises the probability of transition by roughly 33% and being in the highest quintile raises the transition probability by roughly 20%. Column 7 corroborates this finding replacing quintile using a quadratic function of each individual's percentile position within the weekly wage distribution. The coefficients in this specification are precisely estimated. The minimum of the resulting quadratic function occurs between the 50^{th} and 55^{th} percentile. The estimated coefficients indicate that an individual at the 5^{th} percentile is about 35% more likely than an individual at median of the distribution to enter self-employment in the following period and that similarly an individual earning at the 95^{th} percentile is about 28% more likely to enter self-employment. Allowing for a non-linear function of pay does improve the fit of the model modestly and results in a modest decrease in the firm size coefficients. Nonetheless, the estimated coefficients remain highly significant, and tests of their equality are rejected at the $p < .0001$ level.

Table 4b repeats the analysis in Table 4a, using the non-PhD sample. The results of the estimations are nearly identical to those with the PhD sample. We focus only on the key differences. First, the estimated firm size coefficients are less impacted by the addition of additional covariates than they are for the PhD sample; the full set of covariates reduces the magnitude of the firm size coefficients by 17 to 32 % as compared to 31 to 42% in the analysis of PhDs. A second difference is that regional dummies have more explanatory power in this non-PhD sample. The interpretation of these two differences is not entirely clear. If the region and size of firm in which one works and strongly influences

the opportunities that one learns about, it may be the case that non-PhD's decisions are more responsive to opportunities, whereas PhD's decisions to enter self-employment may be more responsive to human capital. Many other explanations, however, are possible. The final difference between the non-PhD and PhD transition analysis is that the non-PhDs who are most likely to enter self employment come from the top of the ability distribution. While a "u"-shaped pattern with respect to pay is observed for both samples, the right side of the "u" is higher for the non-PhD sample (compared with the left side), whereas in the PhD sample the left side of the "u" is higher. The pattern we observe and the differences across the samples raise some interesting questions about the role of ability and pay in self-employment decisions. However, in general the results seem to suggest those leaving small firms for entrepreneurship are disproportionately departing from the tails of the ability distribution.

It is possible, of course, that the observed relationship between firm size and entry into self-employment is simply a function of the increased rate at which employees from smaller businesses change employers. If all employees who separate from a given employer are equally likely to become self-employed, the higher rates of transition we observe when we estimate equation (1) may result from the fact that employees at small firms are simply more likely to leave their jobs than those leaving larger firms. To examine this possibility we estimate a multinomial logit model for those who are not self-employed at time t with the following choices between period t and $t+2$:

1. Remain with current employer in both periods
2. Change jobs, but do not become self-employed
3. Become self employed in $t+2$

If employees changing jobs have a constant rate of entering self-employment, then the ratios of the coefficients on firm size should, roughly speaking, be equal in the equations that predict choices 2 and 3. Table 4c presents the estimates for this model for both samples using the entire set of covariates from column 6 in the probit analyses above. The omitted decision in these estimations is choice 1, remaining with the current employer. The relative magnitude of the estimated coefficients on the small firm dummy variables (size 1-25 and size 26-100) in the self-employment equations (relative to the job change

equations), in both regressions suggest that the size-related differences in transition rates to self employment are not primarily driven by differences in the rates at which employees at firms of different sizes leave their jobs.

Differences in Pay between future entrepreneurs and non-entrepreneurs by Firm Size and type of Self-Employment

In this section we analyze the relationship between ability and self-employment by firm size. In the analysis below, we define small firms as firms with 100 or fewer employees and larger firms as firms with 100 or more employees. We begin by examining the distribution of pay across several categories of respondents. We corroborate this graphical analysis with a series of prospective regressions which allow us to draw more nuanced inferences about the ability of those who become self-employed relative to the ability of those who remain as paid employees.

As Figures 1a and 2a illustrate (for the PhD and non-PhD samples, respectively) the distribution of pay differs significantly in large firms, small firms, and self-employment in ways that are consistent with the prior literature on the relationship between pay and firm size (Rasmusen & Zenger, 1990; Garen, 1985). For both PhDs and non-PhDs, the variance of self employment pay exceeds the variance of pay in small firms which, in turn, exceeds the variance of pay in large firms. Also consistent with the prior literature is the markedly higher average pay among non-PhDs in large firms vs. small firms. However, this difference for the PhD sample is quite modest. Figures 1b and 2b compare the paid employment wages at time t of those who become self-employed by time $t+2$ and those who remain in paid employment. Consistent with the regression analysis in the prior subsection, the future self-employed are more likely to come from the top end or the bottom end of the paid employment distribution than those who remain as paid employees. Figures 1c, 1d, 2c, and 2d, disaggregate this analysis into small firm employees and large firm employees in the two samples, respectively. In the PhD sample, the distributions of pay for future entrepreneurs from both small and large firms have roughly similar means, but greater variances than the distribution of those who remain as paid employees. For the non-PhD

sample, however, the pattern looks somewhat different. Future entrepreneurs from small firms in the non-PhD sample seem to have higher median pay and seem more likely to be drawn from the upper part of the current pay distribution than those who remain in paid employment (see Figure 2c). The distribution of pay among future entrepreneurs from large firms in the non-PhD sample is quite similar to the distribution of pay for "stayers" although they are somewhat more likely to be in the tails of the distribution (see Figure 2d). Overall, these figures provide some evidence that those entering self-employment are more likely to be either "stars" or "slugs" than those remaining in self-employment, although the picture is somewhat more complex.

To examine this relationship in further detail, we estimate "pre-program" regressions of the following form for paid employees:

$$PAY_{it} = \alpha + \beta \boldsymbol{X}_i + \gamma \boldsymbol{Z}_{it} + \delta \boldsymbol{SE}_{i,t+2} + \mu_t + \varepsilon_{it}, \qquad (2)$$

where $\boldsymbol{X}_i$, $\boldsymbol{Z}_{it}$, and μ_t, are, as before, time-invariant individual characteristics, time-varying individual (and individual-job) characteristics, and time dummy variables. PAY_{it} is the individual i's pay at time t, and is measured as the log the weekly wage for individual i at time t. Because we are estimating a continuous variable, $\boldsymbol{Z}_{it}$ contains a richer set of covariates (such as the U.S. state in which the employee works) than in the transition regressions above. The vector $\boldsymbol{SE}_{i,t+2}$ indicates the individual i's future self employment status. In the baseline analysis it is simply a dichotomous variable that equals 1 if the individual i is self employed by period $t+2$ and is 0 otherwise; however, in subsequent analyses we interact this future self-employment status variable with the type of employer at time t, the type of self employment chosen at t+2, and whether the individual is in an R&D track job at time t. In examining the relationship described in equation (2) we estimate an OLS regression and quantile regressions at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles.

Similar to the interpretation of the estimates of equation (1) above, we interpret an individual's pay in the present as a reflection of the ability of the worker in paid employment. Therefore, we interpret the coefficient δ in the OLS regression as indicating whether, on average, those who will become self

employed are "better" employees than those who remain in paid employment, controlling for other major predictors of pay. The interpretation of δ in the quantile regressions is similar. A positive and significant δ in the 90th percentile regression, for example, indicates that the 90th percentile of the pay distribution for the future self-employed is higher than that for those who remain in paid employment, controlling for other factors..

Table 5 presents the results of these estimations on the pooled sample. The pay variable is the log of the top-coded salary divided by the number of weeks worked. All coefficients other than δ are suppressed. In the OLS regressions, robust standard errors are displayed. In the quantile regressions, we present bootstrap standard errors, which are significantly larger than the normal quantile regression standard errors. In analysis series 1, $\boldsymbol{SE}_{i,t+2}$ is a single dichotomous variable indicating whether the individual becomes self-employed in the next period. These analyses show that the median of the pay distribution for the future self-employed is 4.6% higher than those remaining in paid employment, while the differences at the top end and bottom end of the pay distribution are much more pronounced.[11] The 10th percentile of the future entrepreneurs pay distribution is 14.8% lower (significant at $p < .001$) and the 90th percentile is 12.2% higher (significant at $p < .001$), suggesting that entrepreneurs are disproportionately drawn from the high and low ends of the overall pay distribution. In analysis series 2, the dichotomous self-employment variable is interacted with firm size. In this analysis, small firms are characterized as all firms with fewer than 100 employees, and large firms as all firms with more than 100 employees. These analyses suggest the dispersion of pay among those leaving small firm employment to start their own firms is even greater than that among those leaving large firms.

In analysis series 3, the future self employed are split into two groups: those who are incorporated at time *t+2* and those who are not. We note that in the data (but not displayed here) average pay in self-employment is significantly higher for the those who report being incorporated than for those who are not incorporated. Thus incorporation status is one measure of the "success" of the entrepreneur's venture.

[11] Note: we continue to control for firm size in these pay regressions.

Not surprisingly, then, we observe that the pay distribution of future entrepreneurs who incorporate has a higher median and a greater dispersion at upper levels than that of non-entrepreneurs. By contrast, non-incorporated self employed seem to be more concentrated near the low end of the pay distribution, although there is also a disproportionate share near the pay distribution's very high end. Analysis series 4, examines whether these dynamics of the more able migrating to incorporated self employment and the less able to unincorporated self employment is related to firm size. These regressions produce quite pronounced results. Those who leave small firms to become incorporated self-employed are much more likely to be drawn from the top end of the overall pay distribution, whereas those who become non-incorporated self-employed are more likely to be paid less in their jobs. There is limited evidence that the entrepreneurs leaving large firms to become incorporated entrepreneurs are paid more, and some evidence that those who leave large firms to start a non-incorporated business are likely to be higher-paid. We speculate that these employees may be leaving large-firm employment to become independent contractors, perhaps because they lack the knowledge and network accessible through small firms. An alternative possibility is that entrepreneurs from large firms may leave to start businesses that take longer to gestate and begin as non-incorporated activities.

One explanation for these size-related differences in who is drawn into entrepreneurship is that employees in small and large perform differing activities. For instance, we know from Table 3 that employees of large firms are more likely to be solely focused on R&D, while those in small firms are likely to perform a wide range of activities. Perhaps, these differences alter ability in the present job as measured by current pay and therefore determine who enters entrepreneurship from large and small firms. In analysis series 5, we interact the variables from analysis series 2 with dummy variables indicating whether the individual is an "R&D track" employee or not.[12] These regressions show that the increased pay dispersion of future entrepreneurs is coming mainly from non-R&D track employees. By contrast, the pay dispersion of future entrepreneurs who are R&D track employees is not significantly different

[12] Those who are not principally engaged in R&D are likely to be engaged principally in commercial activities and / or management.

than those who remain in paid employment. Thus, non-R&D track employees who become entrepreneurs are more likely to be stars or slugs than their R&D counterparts.

Tables A1 and A2 in the Appendix repeat the analyses of equation (2), performing them separately on the PhD and non-PhD samples. For the PhD sample, the pay measure is constructed using the non-top-coded annual salary. In the main, the results of these separate analyses are highly similar to those in the pooled sample, with a few exceptions, noted below. One difference is that in the PhD sample estimates, the coefficients for the 90th percentile regression are significant only at the $p < .1$ level when future self-employment is interacted with firm size. Another difference is that the coefficient estimates in the median regressions are positive and significant for the non-PhD sample. These positive estimates are highly significant at the median for entrepreneurs coming from small firms, entrepreneurs going to incorporated firms, and those coming from non-R&D track jobs. Thus, on average, relative to the PhD sample, non-PhD entrepreneurs are more likely to come from the top end of the (non-PhD) pay distribution.

5. Small Firm Experience and Entrepreneurial Performance

We now investigate whether the experience of working in a small firm predicts performance in self-employment. For this analysis we pool data from both samples. We focus on two measures of performance: persistence in self-employment, and total pay in the first period of self-employment. We interpret persistence in self-employment as a proxy for survival of the new enterprise, although other interpretations are possible. We interpret pay in the first period of self-employment as a proxy for monetary returns from the enterprise, although clearly this may be a noisy measure.

We parameterize firm size in three ways. First, we create a "continuous" measure of firm size by taking the log of the geometric average of the firm size category endpoints. For the size greater than 5000 category, we arbitrarily set the upper endpoint to 15,000.[13] Second, we employ a single dummy variable

[13] The results that we present below are robust to taking a linear average of the category endpoints and to upper bounds in the top category of 25,000, and 35,000, respectively.

for firms of size 1 – 25. Roughly half of all entrants into self employment come from firms of this size. Finally, we employ multiple dummies for firm size to explore potentially non-linear effects.

Table 6 reports the results of a probit analysis examining the persistence in self employment of new entrepreneurs who vary in their past experience working for small firms. The dependent variable is equal to 1 if the individual reported being self-employed in *t* and in *t + 2* and 0 otherwise. The sample for this analysis includes only those who reported being employed in a firm in *t - 2*. Marginal effects are displayed. Columns 1 through 3 employ the continuous measure of firm size. The simple correlation between persistence and firm size, controlling only for year effects, is negative and significant (see column 1). In column 2 we add control variables, reducing the magnitude of this relationship and its significance. Column 3 further introduces the new entrepreneur's pay at *t - 2*, which we contend proxies for paid employment ability. Adding this control for ability has little impact on the firm size coefficient. Note that the lagged paid employment wage is a positive and significant predictor of entrepreneurial survival, suggesting that abilities are positively correlated across sectors. Including the paid employment wage may thus partially capture differences in entrepreneurial ability. Columns 4 through 6 employ a single dummy variable for new entrepreneurs who worked at firms of size 1 - 25 at *t - 2*, introducing new covariates as above. This analysis suggests that those who come from small firms are roughly 10% more likely to persist in self employment, in absolute terms, than those coming from large firms (note: the sample average is 55%). Columns 7 through 9 corroborate these results, employing more dummy variables to capture the impact of prior employer size. In these regressions, the omitted category is firms of size 101 – 5000. Notably, the sign of the coefficient on the largest size category, 5000+, is positive, although it is not precisely estimated. This raises the possibility that the underlying relationship between persistence in self-employment and firm size may be non-linear.

Table 7 examines the relationship between pay in self-employment and the size of the prior employer. We employ a similar set of covariates as in the estimation of equation (2), dropping only the future self-employment variables. We first employ the continuous measure of firm size in column 1, adding past pay to control for ability in column 2. The predicted coefficient on firm size changes sign

when past pay is added, but in neither case is the coefficient different from zero. The analysis is repeated with the small firm dummy only (columns 3 and 4) and with a set of firm size dummies (columns 5 and 6). In no case are firm size coefficients significantly different from zero. In columns 7 through 9, we repeat the analysis, employing only those whose incomes are in the top half of their respective distributions[14]. The analysis shows that for those entering entrepreneurship from the top half of the distribution, small firm experience is positively related with self-employment pay, controlling for prior pay. Eliminating controls for prior pay, which we examine in unreported regression, strengthens the relationship between past experience and self-employment pay.

While these analyses are informative, they do not allow us to draw causal inferences about the relationship between working at small firm and subsequent performance as an entrepreneur. We attempt to improve our ability to draw a causal inference by using the inverse propensity score weighting method of Hirano and Imbens (2002). This approach allows us to more fully account for potential selection on observable characteristics. While selection on unobservable characteristics may still be important, we include lagged paid employment wage in the propensity score model in an attempt to account for some of this heterogeneity. We report the results of this analysis in Table 8. We consider the "treatment" working in a firm of size 1 – 25 at *t - 2*. In columns 1 – 3, we examine persistence in self employment using a weighted probit model, and in columns 4 – 6 and 7-9 we examine self-employment pay for the entire sample and for the subsample of those coming from the top half of their respective salary distributions at *t - 2*, respectively, using a weighted OLS model.[15] Hirano and Imbens (2002) provide methods for estimating both the average treatment effect and the effect of the treatment on the treated. We present the results of both estimations, together with the baseline model, for the purpose of comparison. Adjusting for the non-random distribution of the treatment, we continue to find significant differences between persistence in self-employment and in self-employment pay (again for the top half of the distribution

[14] For example, we compare the pay of a PhD who transitions to self-employment in 1999 with the pay of other PhDs. in 1997.

[15] We cannot currently perform this weighting using a tobit model.

only) for those having worked in small firms. Although we cannot prove a causal relationship between small firm experience and entrepreneurial experience, neither do we have an immediate basis upon which to rule it out.

6. Conclusion

Our objective has been to examine explanations for the surprisingly large role that small firms play in spawning entrepreneurship. We have empirically examined three categories of explanation: ability sorting, human capital accumulation, and non-pecuniary rewards. Our results provide clear evidence of sorting by ability into self-employment. However, the relationship is by no means a simple one. We find evidence of "two tailed" sorting into self employment. Consistent with prior research suggesting that large firms have difficulty rewarding ability, we find evidence that those who transition from large firms into self employment are disproportionately drawn from the upper end of the pay distribution. However, we see the same pattern that those leaving small firms are also disproportionately drawn from the upper tail of the pay distribution. On the other hand, we also find strong evidence that those entering self-employment from small firms are disproportionately drawn from the lower tail of the distribution, consistent with an opportunity cost and non-pecuniary benefits logic. Those in small firms with low ability and thus correspondingly low pay, presumably choose self-employment because the opportunity cost is low and the non-pecuniary benefits, at least for some, are substantial. Consistent with our explanations for a pattern of two-tailed sorting, we also find that those in the upper, high ability tail are moving into incorporated self-employment, while those in the lower tail who enter self employment more frequently do so without incorporation. We speculate that the greater spawning by small firms may simply reflect the fact that small firms have disproportionately attracted individuals at the two extremes of the distribution. Thus, relative to the distribution of ability in large firms, the distribution of ability in small firms simply possesses a greater preponderance of individuals with incentives to migrate to self employment.

We also find evidence consistent with the idea that employment in small firms provides important access to knowledge, networks, or human capital valuable in founding an entrepreneurial venture. After controlling for the attributes of those who select into small firms and for regional differences, we still find strong evidence that employment in a small firm dramatically enhances the probability of entering self-employment. This small firm effect may partly reflect the fact that small firms push engineers and scientists to perform a much broader set of activities, particularly activities beyond research. Controlling for the breadth of activities an employee performs in their current job does predict movement into self-employment and empirically explains a portion of the small firm effect on entrepreneurial spawning. In this sense, our results are consistent with Lazear's argument that entrepreneurs must be jack of all trades. Employment in small firms appears to play a critical role in providing this broad training valuable in entrepreneurship. We also find evidence that prior employment in small firms enhances the performance of the subsequent entrepreneurial venture.

In summary, our results provide support for all three explanations for small firm spawning that we set out to explore. We find evidence of a pattern of ability sorting into entrepreneurship. We find evidence that employment in small firms provides human capital critical to spawning entrepreneurial ventures. We also find evidence that non-pecuniary rewards play a role in attracting particularly those at the low end of the ability distribution. In future work, we hope to extend the analysis to examine the degree to which high ability wage workers tend to be successful entrepreneurs. Using the semi-parametric panel data selection framework developed by Hamilton and Chib (2002), we can jointly model the choices of scientists to work in entrepreneurial/non-entrepreneurial firms and the outcomes associated with these choices (e.g., wages or self-employment income). This model can be estimated using Markov Chain Monte Carlo simulation methods that allow for correlation in unobserved individual factors that influence choices and outcomes. Positive correlation between unobserved factors affecting returns in each sector would tend to indicate a link between ability in paid employment and success in self-employment.

6. References

Anderson, Pernilla and Eskil Wadensjö. 2006. "Employees Who Become Self-Employed: Do Labor Income and Wages Have an Impact." IZA Discussion Paper No. 1971.

Blanchflower, David G. and A. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16(1): 29-60.

Borjas, George J. and Bronars, Stephen G. "Consumer Discrimination and Self-Employment." *Journal of Political Economy* 108(3): 604- 631.

Braguinsky, Serguey, and Ohyama, Atsushi. 2007. "Where Does Entrepreneurship Pay?" Working Paper, State University of New York, Buffalo.

Brown, Charles and Medoff, James. 1989. "The Employer Size-Wage Effect." *Journal of Political Economy* 97(4): 1027-1059.

Chib, Siddhartha and Barton H. Hamilton. 2002. "Semiparametric Bayes Analysis of Longitudinal Data Treatment Models." *Journal of Econometrics* 110(1): 67-89.

Evans, David S. and Leighton, Linda S. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79(3): 519-535.

Evans, David S. and Jovanovic, Boyan. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97(4): 808-827.

Garen, John E. 1985. "Worker Heterogeneity, Job Screening, and Firm Size." *Journal of Political Economy* 93(4): 715-740.

Gompers, Paul, Lerner, Josh, and David Scharfstein. 2005. "Entrepreneurial Spawning: Public Corporations and the Genesis of New Ventures, 1986 to 1999." *Journal of Finance* 60(2): 577-614.

Hamilton, Barton H. 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment." *Journal of Political Economy* 108(3): 604- 631.

Hellman, Thomas. 2007. "When Do Employees Become Entrepreneurs?" *Management Science* 53(6): 919-933.

Hurst, Erik and Lusardi, Anna Maria. 2004. "Liquidity Constraints, Household Wealth and Entrepreneurship." *Journal of Political Economy* 112(2): 319-347.

Hirano Keisuke and Imbens, Guido W. 2002. "Estimating Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization." Working Paper, the University of Miami and University of California, Berkeley.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50: 53-75.

Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics* 108(3): 551-575.

Kremer, Michael and Maskin, Eric. 1995. "Segregation by Skill and the Rise in Inequality," mimeo.

Lazear, Edward P. 2004. "Balanced Skills and Entrepreneurship." *American Economic Review* 94(2): 208-211

Lazear, Edward P. 2004. "Entrepreneurship." *Journal of Labor Economics* 23(4): 649-680.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9(2): 508-523.

Oi, Walter Y. 1983. "The Fixed Employment Cost of Specialized Labor," in Jack E. Triplett (ed.), *The Measurement of Labor Costs*. Chicago: University of Chicago Press.

Parker, Simon C. 2006. "A Selection-Based Theory of the Transition from Employment to Entrepreneurship: The Role of Employer Size." IZA Discussion Paper #2071.

Rasumusen, Eric and Zenger, Todd R. 1990. "Diseconomies of Scale in Employment Contracts." *Journal of Law, Economics, and Organization* 6(1): 65-92.

Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3: 135-146.

Sauermann, Henry and Cohen, Wesley. "Fire in the Belly? Individual's Motives and Innovative Performance in Startups and Established Firms" Duke University Working Paper.

Stuart, Toby and Ding, Waverly W. 2006. "When Do Scientists Become Entrepreneurs? The Social Structural Antecedents of Commercial Activity in the Academic Life Sciences." *American Journal of Sociology* 112(1): 97-144.

Troske, Kenneth R. 1999. "Evidence on the Employer Size-Wage Premium From Worker-Establishment Matched Data." *Review of Economics and Statistics* 81(1): 15-26.

Velamuri, S. Ramakrishna and Venkataraman, S. 2005. "An Empirical Study of the Transition from Paid Work to Self-Employment." *Journal of Entrepreneurial Finance and Business Ventures* 10(1): 1-16.

Wagner, Joachim. 2004. "Are Young and Small Firms Hothouses for Nascent Entrepreneurs? Evidence from German Micro Data." IZA Discussion Paper #989.

Willis, Robert J. and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87(5):

Zenger, Todd R. 1994. "Explaining Organizational Diseconomies of Scale in R&D: Agency Problems and the Allocation of Engineering Talent, Ideas, and Effort by Firm Size." *Management Science* 40(6): 708-729.

Zucker, Lynn G., Darby, Michael R, and Torero, Maximo. "Labor Mobility from Academe to Commerce." *Journal of Labor Economics* 20(3): 629-660.

Figure 1a. Distributions of Weekly Wages in Large Firms, Small Firms, and Self-Employed, PhD Sample

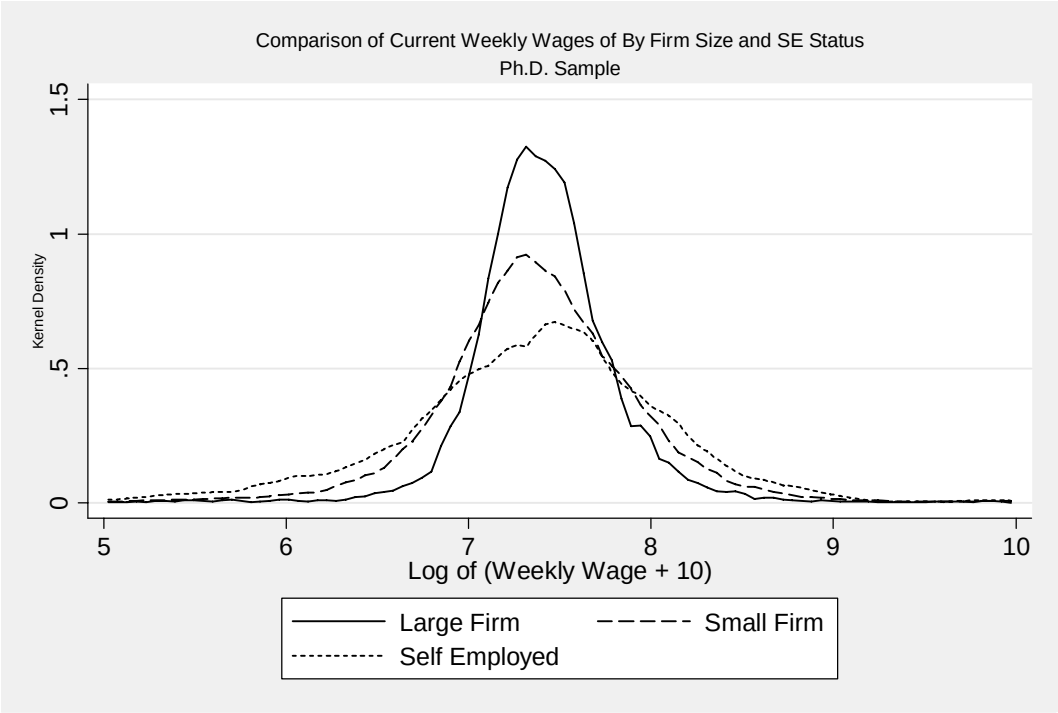

Figure 1b. Distribution of Weekly Wages for Future Self-Employed, Compared with Distribution of Those Remaining in Paid Employment, PhD Sample

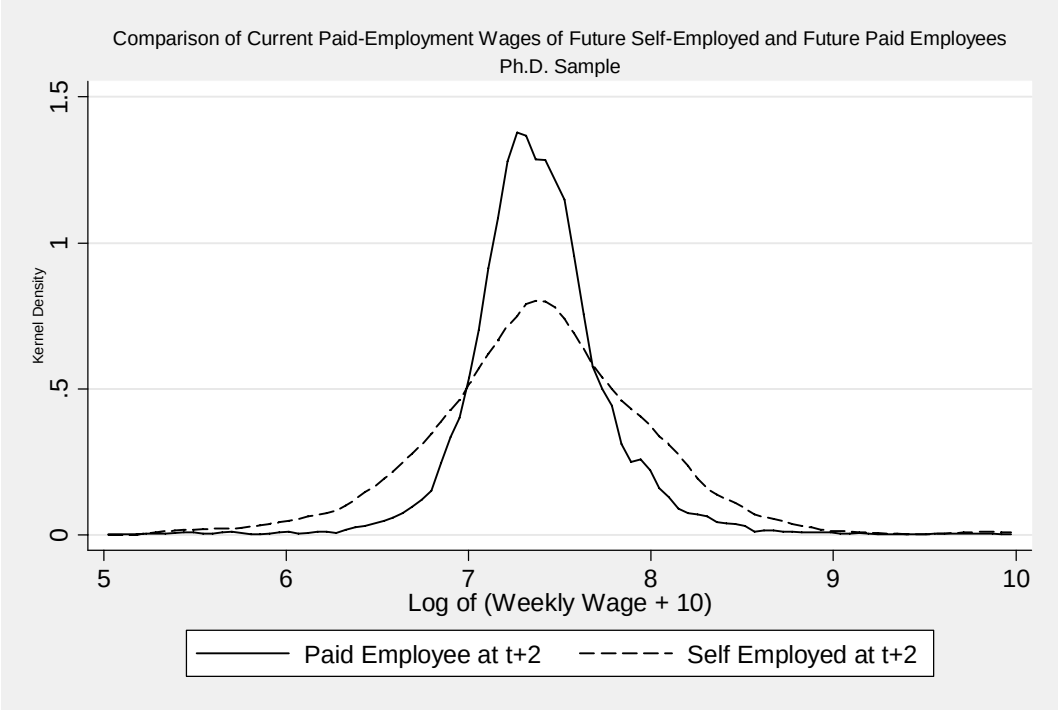

Figure 1c. Distribution of Weekly Wages for Future Self-Employed, Compared with Distribution of Those Remaining in Paid Employment in Small Firms, PhD Sample

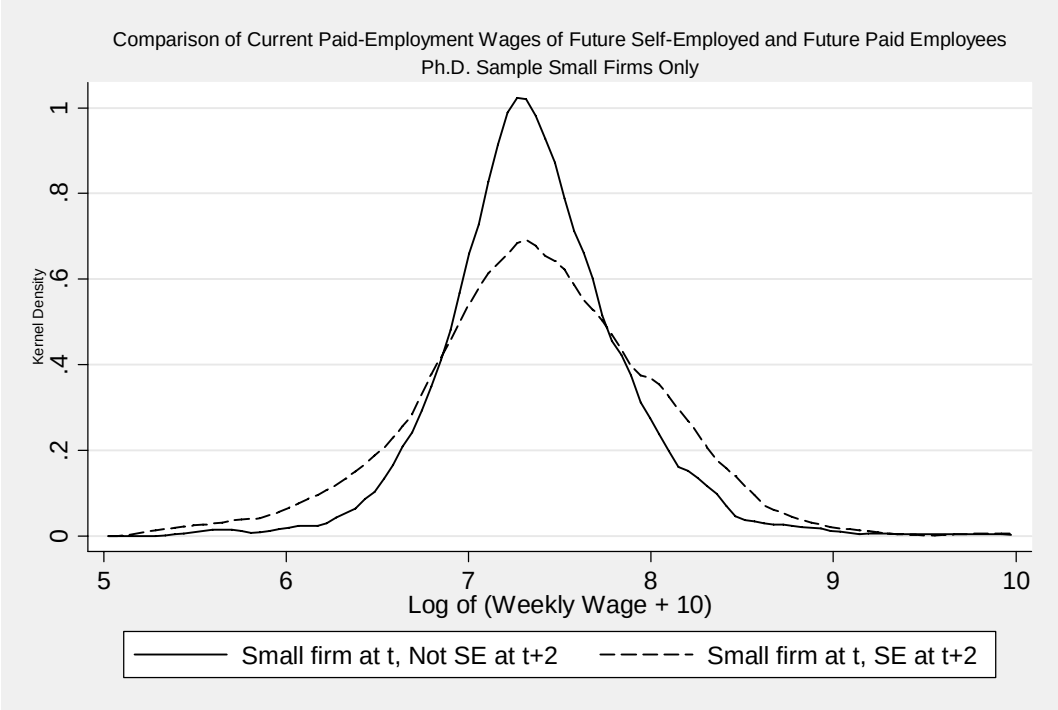

Figure 1d. Distribution of Weekly Wages for Future Self-Employed, Compared with Distribution of Those Remaining in Paid Employment in Large Firms, PhD Sample

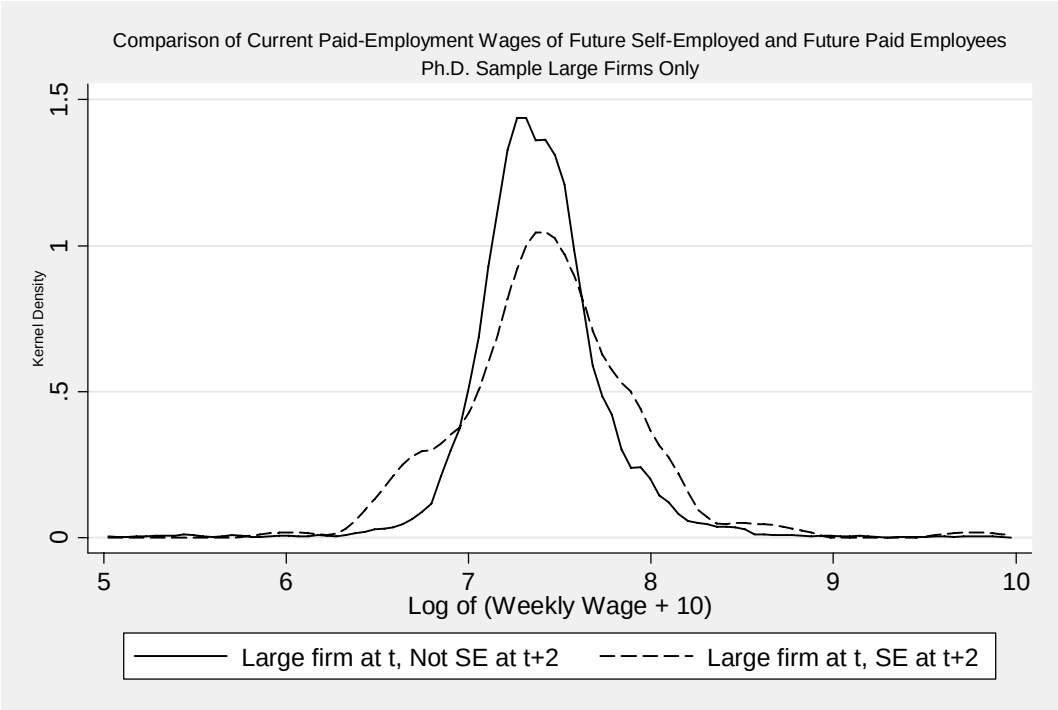

Figure 2a. Distributions of Weekly Wages in Large Firms, Small Firms, and Self-Employed, non-PhD Sample

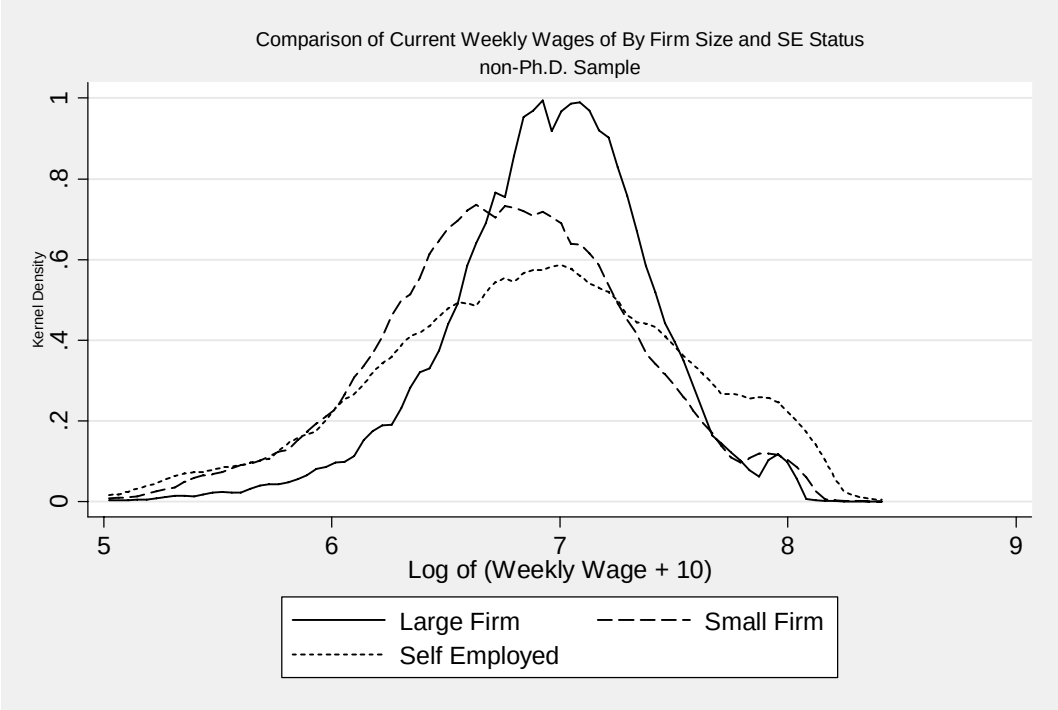

Figure 2b. Distribution of Weekly Wages for Future Self-Employed, Compared with Distribution of Those Remaining in Paid Employment, non-PhD Sample

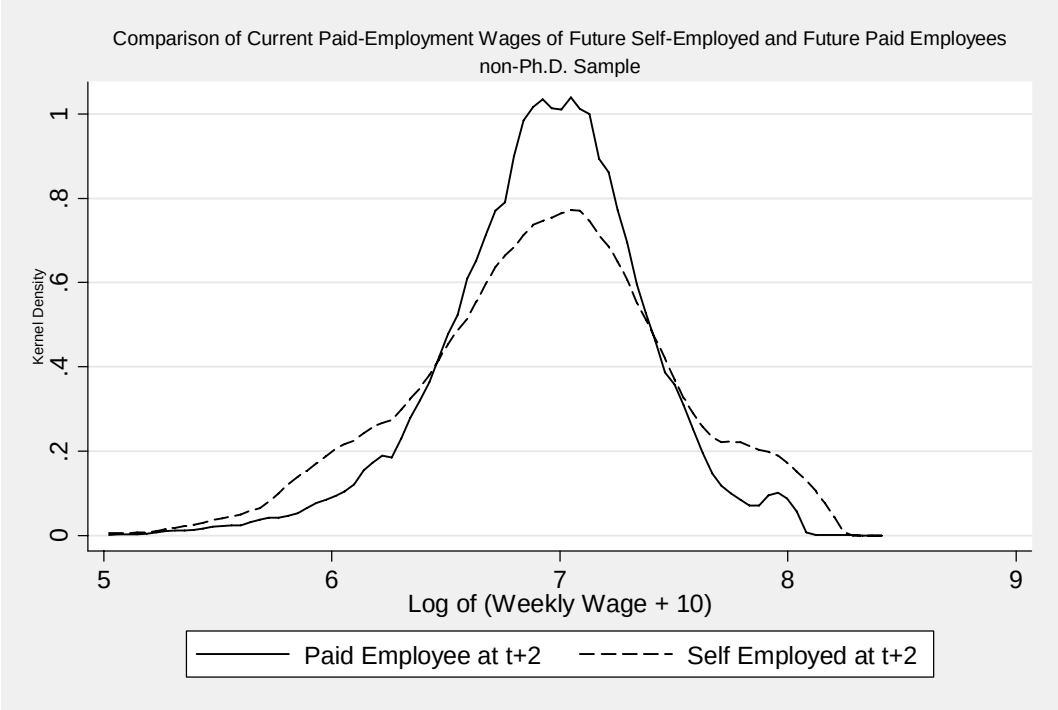

Figure 2c. Distribution of Weekly Wages for Future Self-Employed, Compared with Distribution of Those Remaining in Paid Employment in Small Firms, non-PhD Sample

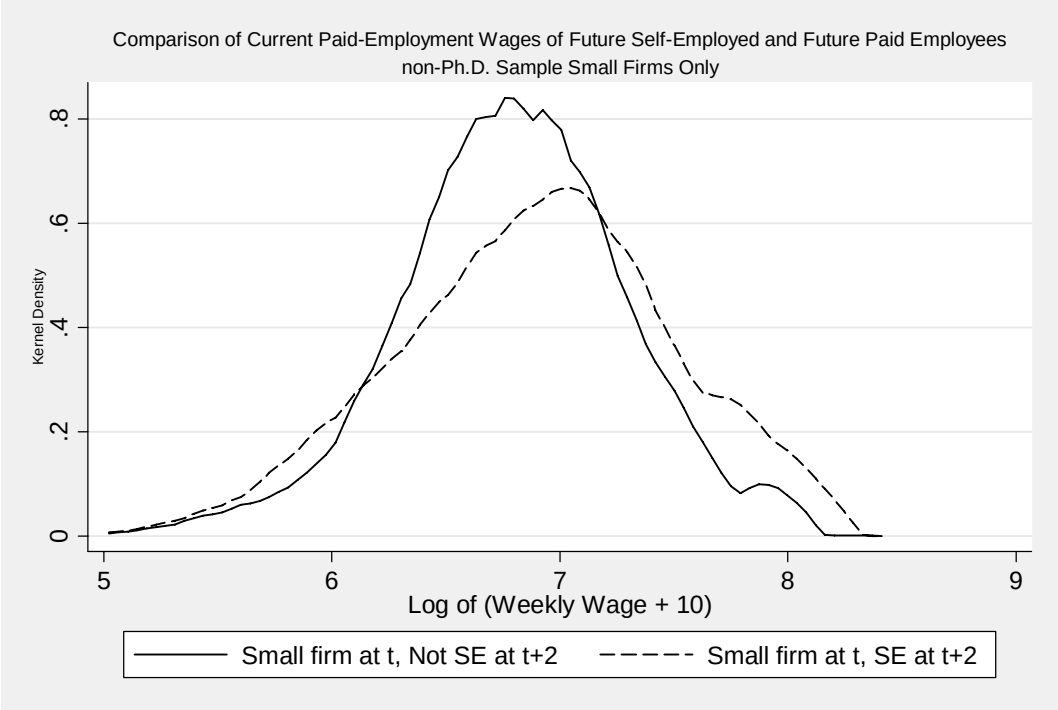

Figure 2d. Distribution of Weekly Wages for Future Self-Employed, Compared with Distribution of Those Remaining in Paid Employment in Large Firms, non-PhD Sample

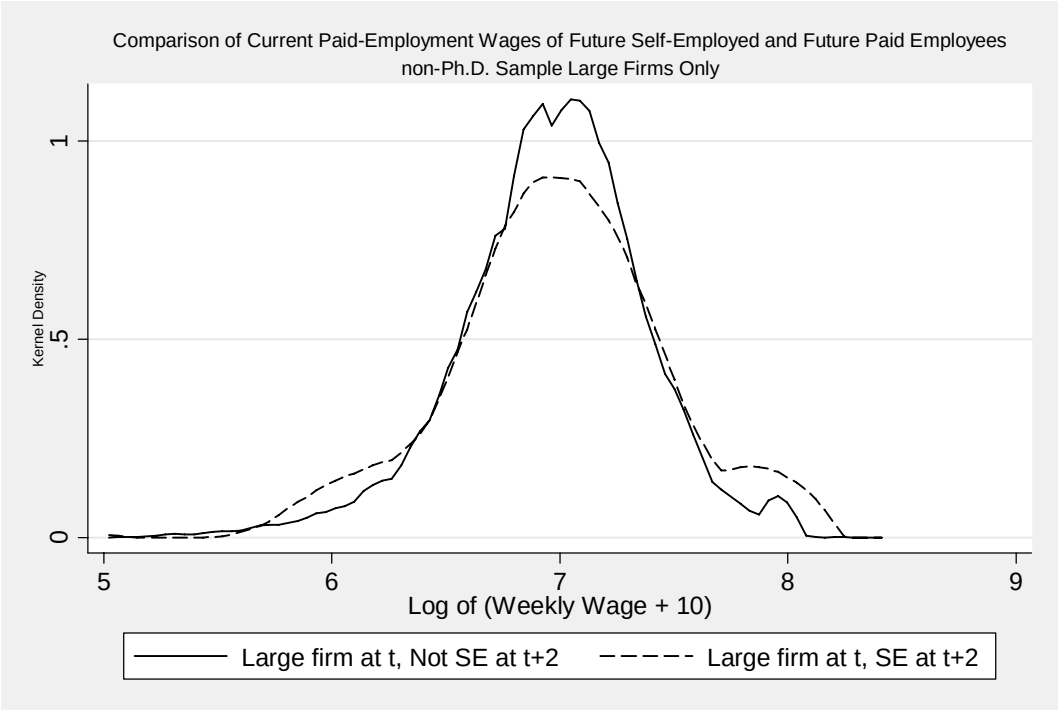

Table 1. Comparison of Job Separation and Transitions into Self-Employment by type of employment in prior survey period. The PhD sample consists of males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. The non-PhD sample consists of males whose responses are included in the SESTAT restricted file in 1995, 1997, 1999, who were at least 22 in 1995 and not more than 65 in 2001, and who did *not* have PhDs. Individuals who were not in the labor force in all relevant periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year.

	Fraction of Employees in:							
	1997		1999		2001		All Years	
Employer Type in Year (t-2)	Turnover	Self-Employed	Turnover	Self-Employed	Turnover	Self-Employed	Turnover	Self-Employed
PhD Sample								
Bus: 1 - 25	28.1%	16.3%	38.5%	22.4%	36.0%	17.9%	34.3%	18.7%
Bus: 26 - 100	21.5%	5.5%	26.6%	6.7%	27.1%	4.2%	25.1%	5.4%
Bus: 101 - 1000	20.0%	2.7%	28.4%	3.2%	28.1%	2.9%	25.3%	2.9%
Bus: 1001 - 5000	15.2%	1.6%	22.4%	2.5%	23.2%	1.8%	20.0%	2.0%
Bus: 5000 +	11.0%	0.8%	13.3%	1.3%	18.0%	0.9%	14.2%	0.9%
Government	7.1%	0.6%	9.7%	0.9%	9.9%	0.6%	8.8%	0.7%
Secondary Ed.	11.2%	1.0%	12.6%	1.4%	9.0%	1.4%	10.8%	1.3%
University / Research Institute	8.4%	0.3%	10.0%	.0.5%	10.6%	0.4%	9.6%	0.4%
Non-PhD Sample								
Bus: 1 - 25	40.2%	15.5%	41.7%	15.6%			40.9%	15.5%
Bus: 26 - 100	32.2%	5.3%	35.7%	5.6%			33.8%	5.4%
Bus: 101 - 1000	27.7%	3.3%	30.5%	2.6%			29.0%	3.0%
Bus: 1001 - 5000	22.0%	1.6%	24.2%	1.8%			23.0%	1.7%
Bus: 5000 +	16.0%	1.0%	17.4%	1.2%	n.a.	n.a.	16.7%	1.1%
Government	8.7%	0.6%	9.3%	0.7%			9.0%	0.6%
Secondary Ed.	12.6%	1.1%	13.0%	0.8%			12.8%	0.9%
University / Research Institute	21.6%	1.7%	21.9%	0.9%			21.8%	1.3%

Notes: n.a. indicates not available

Table 2a. Summary Statistics, PhD sample. The sample includes males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. Individuals in the SDR file who were not in the labor force in all 4 periods are eliminated from the sample. Individuals whose PhDs were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. We include only workers in for-profit enterprise.

	Obs	*Mean*	*Median*	*Std Dev*	*Min*	*Max*
Age	30,583	44.6	44	8.9	23	65
Year	30,583	1997.9	1997	2.24	1995	2001
Years in Current Job	30,583	6.8	4.08	6.9	0	43.7
Hours worked in primary job	30,583	48.2	48	8.4	30	80
Weeks worked in primary job	30,583	51.5	52	1.8	30	52
Salary	30,583	93,442	83,000	61,975	0	999,996
Salary, top-coded as 150,000	30,583	87,146	83,000	32,310	0	150,000
HDF: Computer	30,583	.064	0	.245	0	1
HDF: Life Science	30,583	.220	0	.414	0	1
HDF: Physical Science	30,583	.262	0	.440	0	1
HDF: Social Science	30,583	.116	0	.320	0	1
HDF: Engineering	30,583	.338	0	.473	0	1
White	30,583	.721	1	.449	0	1
Married	30,583	.836	1	.370	0	1
Has spouse who works full time	30,583	.357	0	.479	0	1
Has spouse who works part-time	30,583	.163	0	.370	0	1
Has spouse who does not work	30,583	.315	0	.464	0	1
Children Living in Household	30,583	1.12	1	1.16	0	11
Employer:						
Self-Employed	30,583	.127	0	.332	0	1
Self-Employed, Incorporated	30,583	.054	0	.227	0	1
Self-Employed, Not Inc.	30,583	.073	0	.256	0	1
Business, 1-25 employees	30,583	.098	0	.297	0	1
Business, 26-100 employees	30,583	.081	0	.273	0	1
Business, 101-1000 employees	30,583	.145	0	.352	0	1
Business, 1000 – 5000 emp.	30,583	.119	0	.324	0	1
Business, 5000+ emp.	30,583	.429	0	.495	0	1
Turnover	30,583	.223	0	.416	0	1
Activities on the Job:						
Accounting, Finance, Contracts	30,583	.222	0	.416	0	1
Applied Research	30,583	.637	1	.481	0	1
Basic Research	30,583	.208	0	.406	0	1
Computer Applications	30,583	.416	0	.493	0	1
Development	30,583	.542	1	.498	0	1
Design	30,583	.428	0	.495	0	1
Employee Relations	30,583	.305	0	.460	0	1
Managing or Supervising People	30,583	.571	1	.494	0	1
Other	30,583	.048	0	.213	0	1
Production, Operations, and Maintenance	30,583	.061	0	.240	0	1
Quality or Productivity Management	30,583	.209	0	.407	0	1
Sales, Purchasing, or Marketing	30,583	.239	0	.426	0	1
Professional Services	30,583	.190	0	.392	0	1
Teaching	30,583	.089	0	.286	0	1
Primary Activity is R&D	30,583	.487	0	.500	0	1
Location:						
New England	30,583	.081	0	.273	0	1
Mid Atlantic	30,583	.192	0	.394	0	1
South Atlantic	30,583	.142	0	.349	0	1
East North Central	30,583	.134	0	.340	0	1
West North Central	30,583	.051	0	.220	0	1
East South Central	30,583	.026	0	.159	0	1
West South Central	30,583	.086	0	.281	0	1
Mountain	30,583	.060	0	.237	0	1
Pacific	30,583	.228	0	.419	0	1

Table 2b. Summary Statistics, non-PhD sample. The sample includes males whose responses are included in the SESTAT restricted file in 1995, 1997, and 1999, who did not have PhDs and were at least 22 in 1995 and not more than 65 in 2001. Individuals in the SDR file who were not in the labor force in all 3 periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. We include only workers in for-profit enterprise.

	Obs	*Mean*	*Median*	*Std Dev*	*Min*	*Max*
Age	55,822	37.5	36	10.5	22	65
Year	55,822	1996.9	1997	1.62	1995	1999
Years in Current Job	55,822	5.8	2.8	6.9	0	44.9
Hours worked in primary job	55,822	46.7	45	8.0	30	80
Weeks worked in primary job	55,822	51.6	52	1.6	30	52
Salary	55,822	56,807	52,000	37,967	0	522,000
Salary, top-coded as 150,000	55,822	56,730	52,000	27,481	0	150,000
HD: Bachelor's	55,822	.700	1	.458	0	1
HD: Master's	55,822	.300	0	.458	0	1
HDF: Computer	55,822	.149	0	.356	0	1
HDF: Life Science	55,822	.077	0	.266	0	1
HDF: Physical Science	55,822	.095	0	.293	0	1
HDF: Social Science	55,822	.140	0	.347	0	1
HDF: Engineering	55,822	.539	1	.498	0	1
White	55,822	.770	1	.421	0	1
Married	55,822	.678	1	.468	0	1
Has spouse who works full time	55,822	.311	0	.463	0	1
Has spouse who works part-time	55,822	.135	0	.342	0	1
Has spouse who does not work	55,822	.232	0	.422	0	1
Children Living in Household	55,822	0.90	0	.284	0	17
Employer:						
Self-Employed	55,822	.089	0	.285	0	1
Self-Employed, Incorporated	55,822	.041	0	.199	0	1
Self-Employed, Not Inc.	55,822	.048	0	.213	0	1
Business, 1-25 employees	55,822	.099	0	.299	0	1
Business, 26-100 employees	55,822	.098	0	.298	0	1
Business, 101-1000 employees	55,822	.196	0	.398	0	1
Business, 1000 – 5000 emp.	55,822	.142	0	.350	0	1
Business, 5000+ emp.	55,822	.375	0	.484	0	1
Turnover	55,822	.177	0	.381	0	1
Activities on the Job:						
Accounting, Finance, Contracts	55,822	.287	0	.452	0	1
Applied Research	55,822	.284	0	.451	0	1
Basic Research	55,822	.136	0	.342	0	1
Computer Applications	55,822	.536	1	.499	0	1
Development	55,822	.352	0	.478	0	1
Design	55,822	.467	0	.464	0	1
Employee Relations	55,822	.312	0	.463	0	1
Managing or Supervising People	55,822	.511	1	.500	0	1
Other	55,822	.054	0	.227	0	1
Production, Operations, and Maintenance	55,822	.143	0	.349	0	1
Quality or Productivity Management	55,822	.321	0	.467	0	1
Sales, Purchasing, or Marketing	55,822	.355	0	.478	0	1
Professional Services	55,822	.133	0	.340	0	1
Teaching	55,822	.086	0	.280	0	1
Primary Activity is R&D	55,822	.280	0	.449	0	1
Location:						
New England	55,767	.071	0	.256	0	1
Mid Atlantic	55,767	.141	0	.348	0	1
South Atlantic	55,767	.160	0	.367	0	1
East North Central	55,767	.154	0	.361	0	1
West North Central	55,767	.068	0	.252	0	1
East South Central	55,767	.034	0	.182	0	1
West South Central	55,767	.107	0	.309	0	1
Mountain	55,767	.068	0	.252	0	1
Pacific	55,767	.196	0	.397	0	1

Table 3. Summary Statistics by Self- Employment Status and Firm Size. The PhD sample consists of males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. The non-PhD sample consists of males whose responses are included in the SESTAT restricted file in 1995, 1997, 1999, who were at least 22 in 1995 and not more than 65 in 2001, and who did *not* have PhDs. Individuals who were not in the labor force in all relevant periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. We include only workers in for-profit enterprise.

	Self-Employed	1-25	26-100	100-1000	1000-5000	5000+
PhD Sample						
Age	49.8	46.5	43.8	43.2	43.7	43.5
Year	1997.8	1998.0	1998.1	1998.0	1997.8	1997.9
Years in Current Job	9.3	6.0	4.6	4.7	6	7.4
Hours worked	48.2	49.6	49.6	48.4	47.4	47.7
Weeks Worked	50.4	51.1	51.6	51.6	51.7	51.7
Salary	96,721	93,812	96,044	93,087	92,258	92,347
Salary, top coded	83,690	83,674	89,088	87,217	86,872	88,641
HDF Computer	.038	.065	.062	.073	.059	.070
HDF Life Science	.255	.253	.270	.261	.246	.171
HDF Phys Science	.135	.199	.272	.263	.289	.305
HDF Soc Science	.414	.188	.092	.083	.060	.045
HDF Engineering	.158	.295	.303	.320	.346	.409
White	.822	.766	.730	.698	.702	.761
Supervisor	.466	.663	.697	.639	.595	.583
Commercial Activity Count	1.96	2.09	1.92	1.74	1.49	1.36
Research Activity Count	1.19	1.99	2.22	2.23	2.35	2.57
N	3,881	2,983	2,482	4,456	4,837	13,135
Non-PhD Sample						
Age	43.6	37.1	36.3	36.4	37	37.4
Year	1996.8	1996.8	1996.8	1996.9	1996.8	1996.9
Years in Current Job	8.1	4.7	4.3	4.4	5.3	6.7
Hours worked	50	46.8	46.5	46.5	46.2	46.4
Weeks Worked	50.8	51.5	51.7	51.8	51.8	51.8
Salary	57,405	48,539	52,848	54,154	57,548	60,984
Salary, top coded	57,067	48,539	52,817	54,032	57,479	60,957
HD Bachelor's	.757	.731	.721	.716	.709	.663
HD Master's	.243	.269	.279	.284	.291	.337
HDF Computer	.108	.125	.137	.137	.154	.170
HDF Life Science	.160	.112	.087	.078	.069	.048
HDF Phys Science	.100	.110	.104	.104	.106	.078
HDF Soc Science	.235	.193	.153	.138	.124	.107
HDF Engineering	.398	.459	.518	.541	.545	.595
White	.808	.784	.773	.761	.767	.761
Supervisor	.550	.543	.532	.489	.466	.439
Commercial Activity Count	2.51	2.16	2.00	1.83	1.72	1.57
Research Activity Count	1.28	1.56	1.70	1.79	1.80	1.94
N	4,965	5,522	5,493	10,923	7,976	20,943

Table 4a. Probit Analysis of Transition into Self-Employment by PhD Workers in Paid Employment at For-Profit Firms at Time *t+2* (marginal effects). The sample includes males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. Individuals in the SDR file who were not in the labor force in all 4 periods are eliminated from the sample. Individuals whose PhDs were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. The dependent variable is $SELF\text{-}EMPLOYED_{t+2}$. All regressions are conditional on not being self employed at time *t*. All covariates are at time *t+2*, unless otherwise specified. Standard errors are in brackets.

Column:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bus: $1 - 25_t$	***.2177 [.0139]	***.1622 [.0132]	***.1270 [.0127]	***.1214 [.0131]	***.1262 [.0127]	***.1219 [.0127]	***.1193 [.0125]
Bus: $26 - 100_t$	***.0687 [.0104]	***.0504 [.0090]	***.0452 [.0087]	***.0448 [.0086]	***.0451 [.0087]	***.0438 [.0086]	***.0424 [.0084]
Bus: $101 - 1000_t$	***.0303 [.0063]	***.0205 [.0052]	***.0179 [.0050]	***.0174 [.0050]	***.0178 [.0502]	***.0171 [.0049]	***.0166 [.0048]
Bus: $1001 - 5000_t$	***.0179 [.0060]	**.0134 [.0052]	**.0124 [.0050]	**.0119 [.0050]	**.0124 [.0050]	**.0118 [.0049]	**.0117 [.0049]
Bus: $5000 +_t$	*omitted*	*omitted*	*omitted*	*omitted*	*omitted*	*omitted*	*omitted*
Year = 1999	***.0098 [.0033]	**.0071 [.0028]	**.0069 [.0028]	**.0072 [.0028]	**.0070 [.0028]	**.0067 [.0027]	**.0067 [.0027]
Year = 2001	.0020 [.0028]	.0007 [.0024]	.0011 [.0025]	.0018 [.0025]	.0013 [.0025]	.0011 [.0024]	.0010 [.0024]
HDF: Computer Science		.0041 [.0052]	.0043 [.0054]	.0068 [.0062]	.0045 [.0055]	.0036 [.0052]	.0033 [.0052]
HDF: Life Science		***.0104 [.0036]	*.0079 [.0035]	**.0088 [.0037]	*.0077 [.0036]	*.0071 [.0035]	*.0068 [.0035]
HDF: Social Science		***.0467 [.0086]	***.0250 [.0067]	***.0246 [.0067]	***.0248 [.0067]	***.0239 [.0066]	***.0236 [.0066]
HDF: Engineering Science		.0012 [.0029]	.0004 [.0029]	.0008 [.0029]	.0005 [.0029]	.0005 [.0028]	.0004 [.0028]
Age		.0012 [.0007]	.0005 [.0008]	.0002 [.0008]	.0005 [.0008]	.0003 [.0008]	.0003 [.0008]
Age Squared * 100		.0004 [.0014]	.0003 [.0014]	.0008 [.0014]	.0003 [.0014]	.0004 [.0014]	.0004 [.0014]
Job $Tenure_t$		.0000 [.0004]	-.0001 [.0004]	-.0000 [.0004]	-.0000 [.0004]	.0000 [.0004]	-.0000 [.0004]
Job $Tenure\ Squared_t$* 100		.0006 [.0018]	.0006 [.0018]	.0003 [.0017]	.0007 [.0017]	.0004 [.0016]	.0005 [.0015]
White		-.0046 [.0028]	-.0045 [.0031]	-.0047 [.0027]	-.0045 [.0026]	-.0045 [.0026]	-.0047 [.0026]
Spouse works full time		.0025 [.0034]	.0028 [.0032]	.0029 [.0032]	.0029 [.0032]	.0029 [.0032]	.0031 [.0032]
Spouse works part time		-.0003 [.0038]	-.0001 [.0036]	.0000 [.0036]	.0001 [.0036]	.0000 [.0036]	.0000 [.0036]
Spouse does not work		-.0056 [.0033]	-.0048 [.0031]	-.0046 [.0032]	-.0046 [.0032]	-.0050 [.0031]	-.0051 [.0031]
Children in the Household		.0014 [.0010]	.0013 [.0009]	.0013 [.0009]	.0013 [.0009]	.0014 [.0009]	.0014 [.0009]
$Regional\ Dummies_t$		Y	Y	Y	Y	Y	Y
$Commercial\ Activity\ Dummies_t$		N	***Y	***Y	***Y	***Y	***Y
$Research\ Activity\ Dummies_t$		N	*Y	*Y	*Y	Y	Y
Predicted PE log weekly wage				-.0034 [.0036]			
Predicted SE log weekly wage				.0037 [.0030]			
$Log\ Weekly\ Wage_t$					-.0009 [.0017]		
$Weekly\ Wage\ Quintile_t == 1$						*.0118 [.0057]	
$Weekly\ Wage\ Quintile_t == 2$						.0022 [.0041]	
$Weekly\ Wage\ Quintile_t == 4$						.0015 [.0031]	
$Weekly\ Wage\ Quintile_t == 5$						*.0075 [.0035]	
$Weekly\ Wage_t$ Percentile / 100							***-.0597 [.0151]
$Weekly\ Wage_t$ Pctle. / 100 Squared							***.0566 [.0134]
Obs P.	.0357	.0357	.0357	.0357	.0357	.0357	.0357
N	13,759	13,759	13,759	13,759	13,759	13,759	13,759
Pseudo-R2	.1713	.2074	.2313	.2317	.2313	.2339	.2356

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test)

Note: For firm size category variables, the omitted variable is Size > 5000. The omitted variable for high degree field (HDF) is physical science. Non-married is the omitted category for marital status.

Table 4b. Probit Analysis of Transition into Self-Employment by non-PhD Workers in Paid Employment at For-Profit Firms at Time *t+2* (marginal effects). The sample includes males whose responses are included in the SESTAT restricted file in 1995, 1997, 1999, who were at least 22 in 1995 and not more than 65 in 2001, and who did *not* have PhDs. Individuals in the SESTAT file who were not in the labor force in all 4 periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. The dependent variable is $SELF\text{-}EMPLOYED_{t+2}$. All regressions are conditional on not being self employed at time *t*. All covariates are at time *t+2*, unless otherwise specified. Standard errors are in brackets.

Column:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bus: $1 - 25_t$	***.1870 [.0139]	***.1740 [.0108]	***.1507 [.0104]	***.1570 [.0119]	***.1534 [.0106]	***.1557 [.0108]	***.1519 [.0107]
Bus: $26 - 100_t$	***.0661 [.0076]	***.0630 [.0075]	***.0537 [.0069]	***.0548 [.0070]	***.0545 [.0070]	***.0554 [.0071]	***.0545 [.0070]
Bus: $101 - 1000_t$	***.0288 [.0044]	***.0272 [.0042]	***.0240 [.0040]	***.0249 [.0042]	***.0243 [.0040]	***.0249 [.0041]	***.0245 [.0041]
Bus: $1001 - 5000_t$	***.0103 [.0043]	**.0092 [.0041]	*.0072 [.0038]	.0068 [.0039]	*.0074 [.0038]	*.0076 [.0038]	*.0075 [.0038]
Bus: $5000 +_t$	*omitted*	*omitted*	*omitted*	*Omitted*	*omitted*	*omitted*	*omitted*
Year = 1999	.0002 [.0019]	.0002 [.0018]	.0001 [.0017]	-.0008 [.0021]	.0001 [.0017]	.0002 [.0017]	.0003 [.0017]
HD: Master's		*-.0045 [.0019]	-.0030 [.0019]	-.0031 [.0020]	-.0031 [.0020]	-.0035 [.0019]	-.0036 [.0019]
HDF: Computer Science		.0019 [.0039]	.0027 [.0040]	.0023 [.0040]	.0026 [.0039]	.0022 [.0038]	.0020 [.0038]
HDF: Life Science		*.0102 [.0054]	.0058 [.0048]	.0062 [.0049]	.0059 [.0048]	.0060 [.0048]	.0059 [.0048]
HDF: Social Science		*.0083 [.0044]	.0021 [.0038]	.0017 [.0039]	.0021 [.0038]	.0020 [.0038]	.0018 [.0038]
HDF: Engineering Science		-.0029 [.0031]	-.0026 [.0031]	-.0028 [.0032]	-.0026 [.0031]	-.0027 [.0031]	-.0027 [.0031]
Age		.0002 [.0004]	.0000 [.0004]	-.0002 [.0005]	.0000 [.0004]	-.0001 [.0004]	-.0001 [.0004]
Age Squared * 100		.0015 [.0009]	.0015 [.0009]	.0019 [.0010]	.0015 [.0009]	.0016 [.0009]	.0016 [.0009]
Job $Tenure_t$		.0005 [.0004]	.0004 [.0004]	.0004 [.0004]	.0004 [.0004]	.0004 [.0004]	.0004 [.0004]
Job $Tenure_t$ Squared* 100		-.0020 [.0014]	-.0015 [.0014]	-.0014 [.0014]	-.0015 [.0014]	-.0016 [.0014]	-.0017 [.0014]
White		-.0024 [.0024]	-.0028 [.0026]	-.0033 [.0024]	-.0029 [.0024]	-.0032 [.0024]	-.0033 [.0024]
Spouse works full time		.0033 [.0027]	.0029 [.0026]	.0027 [.0026]	.0028 [.0026]	.0029 [.0026]	.0031 [.0026]
Spouse works part time		.0003 [.0032]	-.0003 [.0031]	-.0007 [.0031]	-.0003 [.0031]	-.0008 [.0031]	-.0004 [.0031]
Spouse does not work		-.0034 [.0028]	-.0039 [.0027]	-.0043 [.0027]	-.0039 [.0027]	-.0045 [.0027]	-.0043 [.0027]
Children in the Household		.0016 [.0008]	.0015 [.0008]	.0015 [.0008]	.0015 [.0008]	.0014 [.0008]	.0014 [.0008]
Regional $Dummies_t$		*Y	*Y	*Y	*Y	*Y	Y
Commercial Activity $Dummies_t$		N	***Y	***Y	***Y	***Y	***Y
Research Activity $Dummies_t$		N	*Y	*Y	*Y	*Y	*Y
Predicted PE log weekly wage				.0085 [.0090]			
Predicted SE log weekly wage				-.0051 [.0107]			
Log Weekly $Wage_t$					.0026 [.0017]		
Weekly Wage $Quintilet_t$ == 1						.0008 [.0034]	
Weekly Wage $Quintile_t$ == 2						-.0051 [.0026]	
Weekly Wage $Quintile_t$ == 4						.0009 [.0027]	
Weekly Wage $Quintile_t$ == 5						*.0063 [.0029]	
Weekly $Wage_t$ Percentile / 100							*-.0304 [.0136]
Weekly $Wage_t$ Pctle. / 100 Squared							***.0388 [.0120]
Obs P.	.0356	.0356	.0356	.0356	.0356	.0356	.0356
N	23,883	23,869	23,869	23,869	23,869	23,869	23,869
Pseudo-R2	.1246	.1447	.1545	.1548	.1549	.1569	.1574

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test)

Note: For firm size category variables, the omitted variable is Size > 5000. The omitted variable for high degree field (HDF) is physical science, and the Master's degree is the omitted category for highest degree (HD). Non-married is the omitted category for marital status.

Table 4c. Multinomial Logit Analysis of Likelihood of Entering Self-Employment or Changing Jobs Among PhD Workers in Paid Employment at For-Profit Firms. The PhD sample includes males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. The non-PhD sample includes males whose responses are included in the SESTAT restricted file in 1995, 1997, 1999, who were at least 22 in 1995 and not more than 65 in 2001, and who did *not* have PhDs. Individuals who were not in the labor force in all relevant periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. All regressions are conditional on not being self employed at time *t-2*. All covariates are at time *t*, unless otherwise specified. Standard errors are in brackets.

	PhD Sample		Non-PhD Sample	
Choice:	Change Jobs, Not Self Employed	Self-Employed	Change Jobs, Not Self Employed	Self-Employed
Column:	(1a)	(1b)	(2a)	(2b)
Bus: $1 - 25_t$	***.3215 [.0871]	***2.5668 [.1645]	***.6681 [.0581]	***2.8083 [.1289]
Bus: $26 - 100_t$	***.3499 [.0848]	***1.6042 [.1901]	***.6938 [.0555]	***1.6608 [.1367]
Bus: $101 - 1000_t$	***.4920 [.0657]	***.9585 [.1892]	***.5278 [.0444]	***1.0875 [.1331]
Bus: $1001 - 5000_t$	***.2893 [.0709]	***.6949 [.2121]	***.3167 [.0502]	*.3964 [.1663]
Bus: $5000 +_t$	*omitted*	*omitted*	*Omitted*	*omitted*
Year = 1999	***.2332 [.0618]	**.3931 [.1260]	*.0751 [.0335]	.0025 [.0753]
Year = 2001	***.3910 [.0572]	.1917 [.1279]		
HD: Bachelor's			*-.1006 [.0039]	.1306 [.0886]
HDF: Computer Science	***.4306 [.0980]	.3421 [.2506]	***.2337 [.0716]	.1421 [.1576]
HDF: Life Science	.1139 [.0724]	**.4296 [.1545]	-.0891 [.0873]	.1921 [.1638]
HDF: Social Science	*.2334 [.1033]	***.8673 [.1697]	.1289 [.0756]	.1104 [.1478]
HDF: Engineering Science	**.1900 [.0633]	.0952 [.1546]	-.0597 [.0626]	-.1582 [.1309]
Age	**.0359 [.0182]	.0518 [.0407]	*-.0200 [.0088]	-.0125 [.0184]
Age Squared * 100	***-.1036 [.0366]	-.0416 [.0743]	-.0078 [.0209]	*-.0805 [.0397]
Job Tenure_t	***-.1125 [.0115]	-.0174 [.0218]	***-.1333 [.0082]	-.0072 [.0151]
Job Tenure_t Squared* 100	***.2732 [.0504]	.0822 [.0878]	***.3044 [.0356]	.0061 [.0570]
White	**-.1730 [.0547]	*-.2577 [.1248]	**-.1182 [.0421]	-.1459 [.0956]
Spouse works full time	.0059 [.0756]	.1162 [.1557]	.0029 [.0459]	.0838 [.1060]
Spouse works part time	*-.1850 [.0926]	-.0458 [.1820]	-.0789 [.0625]	-.0383 [.1348]
Spouse does not work	.0163 [.0813]	-.2454 [.1776]	.0417 [.0541]	-.1962 [.1254]
Children in the Household	-.0214 [.0243]	.0427 [.0474]	*-.0378 [.0173]	.0581 [.0336]
Regional Dummies_t	***Y	Y	***Y	*Y
Commercial Activity Dummies_t	Y	***Y	*Y	***Y
Research Activity Dummies_t	**Y	**Y	**Y	*Y
Predicted PE log weekly wage		-.2439 [.1728]		.3671 [.3835]
Predicted SE log weekly wage		.1973 [.1505]		-.2392 [.4603]
N	13,759		23,869	
Pseudo-R2	.1032		.0991	

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test)

Table 5. Prospective analysis of current salary on likelihood of moving into self-employment in subsequent periods for pooled sample. The PhD subsample includes males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. The non-PhD subsample includes males whose responses are included in the SESTAT restricted file in 1995, 1997, 1999, who were at least 22 in 1995 and not more than 65 in 1999, and who did *not* have PhDs. These two samples are pooled together for the analysis in the table below. Individuals who were not in the labor force in all relevant periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. The dependent variable is log of the weekly wage at time t. All regressions are conditional on not being self employed at time *t*. All covariates are at time *t*, unless otherwise specified. For the OLS regression, robust standard errors are presented (in brackets). For the quantile regressions, bootstrap standard errors with 100 repetitions are presented. N = 37,607.

Analysis Series	Key Coefficients	OLS	Quantile: 10%	Quantile: 25%	Quantile: 50%	Quantile: 75%	Quantile: 90%
(1)	$SELF\text{-}EMPLOYED_{t+2}$	-.0186 [0196]	***-.1485 [.0355]	*-.0431 [.0207]	*.0457 [.0158]	***.0779 [.0165]	***.1215 [.0168]
(2)	$SMALLFIRM_t$ *$SELF\text{-}EMPLOYED_{t+2}$	-.0421 [.0286]	**-.2009 [.0665]	*-.0799 [.0348]	*.0637 [.0227]	***.1188 [.0259]	***.1255 [.0257]
	$LARGEFIRM_t$ *$SELF\text{-}EMPLOYED_{t+2}$	.0190 [.0219]	*-.0833 [.0385]	-.0192 [.0182]	.0283 [.0219]	***.0609 [.0171]	**.0858 [.0297]
(3)	SE_INC_{t+2}	.0308 [.0249]	-.0573 [.0373]	.0089 [.0226]	***.0679 [.0169]	***.1176 [.0215]	***.1326 [.0290]
	$SE\_NOT\_INC_{t+2}$	**-.0881 [.0312]	***-.3385 [.0721]	***-.1488 [.0390]	.0013 [.0335]	.0189 [.0268]	***.0944 [.0237]
(4)	$SMALLFIRM_t$ *SE_INC_{t+2}	.0343 [.0351]	†-.0906 [.0539]	.0242 [.0325]	***.0936 [.0231]	***.1808 [.0398]	***.1761 [.0306]
	$SMALLFIRM_t$*$SE\_NOT\_INC_{t+2}$	***-.1570 [.0466]	***-.4142 [.0866]	***-.2178 [.0423]	-.0287 [.0376]	-.0185 [.0292]	.0631 [.0445]
	$LARGEFIRM_t$*SE_INC_{t+2}	.0190 [.0298]	-.0123 [.0457]	.0002 [.0226]	.0168 [.0280]	*.0514 [.0213]	.0211 [.0463]
	$LARGEFIRM_t$*$SE\_NOT\_INC_{t+2}$	.0189 [.0322]	†-.1450 [.0793]	-.0573 [.0397]	.0389 [.0365]	*.0836 [.0407]	***.1527 [.0426]
(5)	$SMALLFIRM_t$ *$RNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	-.0835 [.0618]	-.1381 [.1211]	-.0215 [.0794]	-.0209 [.0472]	.0583 [.0499]	.0424 [.0453]
	$SMALLFIRM_t$ *$NOTRNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	-.0307 [.0321]	**-.2219 [.0741]	-.0845 [.0296]	***.0751 [.0219]	***.1279 [.0295]	***.1571 [.0284]
	$LARGEFIRM_t$ *$RNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	.0414 [.0256]	.0272 [.0450]	.0241 [.0290]	.0200 [.0266]	.0287 [.0311]	.0274 [.0372]
	$LARGEFIRM_t$ *$NOTRNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	.0119 [.0276]	**-.1161 [.0401]	†-.0417 [.0245]	.0367 [.0315]	**.0721 [.0253]	**.1296 [.0428]

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test); † = significant at $p \leq 0.1$ (two-sided test)

Notes: Variables in wage equation not displayed above include controls for the state in which the employee works, the employee's highest degree, the field of the highest degree, age and age squared, job tenure and job tenure squared, race dummy variables, small firm dummy variable, and year dummies.

Table 6. Probit analysis of Persistence in Self-Employment among newly self-employed by size of previous employer (marginal effects). The sample consist of all members of the pooled sample who moved from employment in a for-profit business to self-employment and who responded to the survey in the subsequent period. The dependent variable is a dichotomous variable, equal to 1 if the individual transitioning into self employment was self-employed at both time *t* and time *t+2* and equal to 0 if the individual was self-employed at time *t* but reported working in another job at time *t+2*. Firm size and salary variables refer to the individual's employer immediately prior to transitioning into self-employment and are measured at *t-2*. Standard errors are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Firm Size$_{t-2}$	*-.0164	*-.0114	†-.0119						
	[.0065]	[.0068]	[.0069]						
Firm Size: 1-25 $_{t-2}$				***.1262	**.0995	*.1023	*.1141	*.0947	*.0972
				[.0369]	[.0382]	[.0385]	[.0463]	[.0474]	[.0476]
Firm Size: 26-100 $_{t-2}$							-.0584	-.0487	-.0484
							[.0632]	[.0639]	[.0639]
Firm Size: 5000+ $_{t-2}$							.0186	.0397	.0387
							[.0662]	[.0559]	[.0669]
Year = 1999	.0326	.0613	.0603	.0299	.0617	.0605	.0281	.0614	.0603
	[.0451]	[.0613]	[.0614]	[.0452]	[.0614]	[.0614]	[.0452]	[.0615]	[.0615]
Year = 2001	***.2572	***.2782	***.2779	***.2568	***.2800	***.2798	***.2569	***.2808	***.2805
	[.0429]	[.0542]	[.0542]	[.0429]	[.0542]	[.0541]	[.0430]	[.0542]	[.0542]
HD: Master's		.0456	.0447		.0503	.0492		.0523	.0512
		[.0649]	[.0649]		[.0649]	[.0649]		[.0614]	[.0650]
HD: PhD		-.0567	-.0611		-.0580	-.0633		-.0594	-.0643
		[.0606]	[.0612]		[.0606]	[.0605]		[.0607]	[.0614]
Age		*.0217	*.0215		*.0210	*.0208		†.0205	†.0203
		[.0106]	[.0106]		[.0106]	[.0106]		[.0106]	[.0106]
Age Squared * 100		-.0327	-.0324		-.0321	-.0318		-.0311	-.0307
		[.0215]	[.0215]		[.0212]	[.0212]		[.0212]	[.0212]
White		***.1801	***.1794		***.1764	***.1757		***.1782	***.1775
		[.0465]	[.0466]		[.0467]	[.0464]		[.0467]	[.0467]
Log Salary$_{t-2}$			.0063			.0101			.0095
			[.0183]			[.0184]			[.0184]
N	735	735	735	735	735	735	735	735	735
Observed P.	.5497	.5497	.5497	.5497	.5497	.5497	.5497	.5497	.5497
Pseudo R^2	.0375	.0629	.0631	.0427	.0666	.0669	.0440	.0681	.0683

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test); † = significant at $p \leq 0.1$ (two-sided test)

Table 7. Tobit analysis of first period self-employment earnings by size of previous employer. The sample consists of all members of the pooled sample who moved from employment in a for-profit business to self-employment. The dependent variable is the log of the salary reported in the first period of self employment, top-coded at 150,000. Firm size and salary variables refer to the individual's employer immediately prior to transitioning into self-employment and are measured at *t-2*. State dummy variables (e.g., AK, AR, AZ, etc.), race dummy variables (African-American, Asian, and Hispanic), and dummy variables for the field of the individual's highest degree (computer science, physical science, life science, social science, and engineering) are included in the regressions below but are not reported. Standard errors are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Firm Size$_{t-2}$	.0120	-.0081					*-.0288		
	[.0108]	[.0104]					[.0124]		
Firm Size: 1-25 $_{t-2}$			-.0510	.0464	-.0676	.0287		*.1718	†.1698
			[.0614]	[.0592]	[.0749]	[.0719]		[.0735]	[.0871]
Firm Size: 26-100 $_{t-2}$					-.0524	-.0282			.0154
					[.1000]	[.0956]			[.1164]
Firm Size: 5000+ $_{t-2}$					-.0085	-.0438			-.0243
					[.1056]	[.1009]			[.1159]
Year = 1999	.0274	-.0611	.0275	-.0615	.0279	-.0603	-.0083	-.0080	-.0077
	[.0644]	[.0620]	[.0632]	[.0620]	[.0654]	[.0620]	[.0761]	[.0760]	[.0761]
Year = 2001	.0326	-.0336	.0311	-.0339	.0315	-.0321	.2005	.1981	.1985
	[.1116]	[.1069]	[.1116]	[.1069]	[.1116]	[.1069]	[.1382]	[.1381]	[.1381]
HD: Master's	.0862	.0784	.0865	.0792	.0876	.0796	.0609	.0550	.0609
	[.0880]	[.0841]	[.0880]	[.0841]	[.0881]	[.0841]	[.1011]	[.1013]	[.1012]
HD: PhD	***.4987	*.3415	***.5007	*.3419	.5027	*.3475	-.2210	-.22287	-.2262
	[.1537]	[.1472]	[.1537]	[.1472]	[.1539]	[.1475]	[.1995]	[.1996]	[.2000]
Age	*.0347	.0132	†.0345	.0132	†.0341	.0127	-.0110	-.0109	-.0110
	[.0176]	[.0169]	[.0176]	[.0169]	[.0176]	[.0169]	[.0227]	[.0232]	[.0229]
Age Squared * 100	*-.0743	-.0380	*-.0741	-.0381	*-.0734	-.0372	.0014	-.0011	.0013
	[.0328]	[.0315]	[.0329]	[.0315]	[.0329]	[.0316]	[.0418]	[.0416]	[.0418]
Log Salary$_{t-2}$		***.4251		***.4244		***.4247	***1.0171	***1.0155	***1.0148
		[.0340]		[.0339]		[.0339]	[.1054]	[.1054]	[.1054]
N	1520	1520	1520	1520	1520	1520	987	987	987
Log Likelihood	-2219.69	-2144.57	-2219.92	-2144.57	-2219.78	-2144.46	-1317.35	-1317.31	-1317.26
Pseudo R^2	.0271	.0600	.0270	.0600	.0270	.0600	.0719	.0719	.0720

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test); † = significant at $p \leq 0.1$ (two-sided test)

Table 8. Propensity Score Adjustment to Performance Analyses. The sample consists of all members of the pooled sample who moved from employment in a for-profit business to self-employment. In columns 1-3, the dependent variable is dichotomous, equal to 1 if the individual transitioning into self employment was self-employed at both time *t* and time *t+2* and equal to 0 if the individual was self-employed at time *t* but reported working in another job at time *t+2*. Firm size and salary variables refer to the individual's employer immediately prior to transitioning into self-employment and are measured at *t-2*. In columns 4-9, the dependent variable is the log of the salary reported in the first period of self employment, top-coded at 150,000. Firm size and salary variables refer to the individual's employer immediately prior to transitioning into self-employment and are measured at *t-2*. Region dummy variables, race dummy variables (African-American, Asian, and Hispanic), and dummy variables for the field of the individual's highest degree (computer science, physical science, life science, social science, and engineering) are included in these regressions but are not reported. Standard errors are in brackets.

Specification:	Probit			OLS					
Dependent Variable:	Persistence in Self-Employment			Self Employment Earnings					
Weighting Adjustment:	None	Average Treatment	Treatment on Treated	None	Average Treatment	Treatment on Treated	None	Average Treatment	Treatment on Treated
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Firm Size: 1-25 $_{t-2}$	*.1023	†.0775	†.0600	.0450	.0417	.0457	*.1451	†.1224	*.1731
	[.0385]	[.0406]	[.0436]	[.0531]	[.0548]	[.0632]	[.0641]	[.0679]	[.0739]
Year = 1999	.0605	.0579	.0606	-.0551	-.0614	-.1152	-.0125	-.0057	-.0098
	[.0614]	[.0674]	[.0700]	[.0557]	[.0581]	[.0668]	[.0671]	[.0713]	[.0761]
Year = 2001	***.2798	***.2838	***.2973	-.0440	-.0029	-.0721	.0761	.1299	.1342
	[.0541]	[.0579]	[.0563]	[.0930]	[.1040]	[.1030]	[.1382]	[.1430]	[.1348]
HD: Master's	.0492	.0163	.0744	.0670	†.1305	.1002	.0838	†.1238	.0665
	[.0649]	[.0695]	[.0731]	[.0772]	[.0706]	[.0706]	[.0868]	[.0688]	[.0781]
HD: PhD	-.0633	-.0618	-.0580	*.1894	*.1747	*.2548	-.0375	-.0202	-.0212
	[.0605]	[.0667]	[.0677]	[.0706]	[.0841]	[.0915]	[.0808]	[.1026]	[.1052]
Age	*.0208	**.0331	**.0322	†.0228	†.0213	†.0220	.0002	0071	0124
	[.0106]	[.0119]	[.0133]	[.0120]	[.0116]	[.0116]	[.0169]	[.0156]	[.0187]
Age Squared * 100	-.0318	*-.0557	*-.0564	*-.0536	*-.0493	*-.0496	-.0131	-.0256	-.0351
	[.0212]	[.0239]	[.0270]	[.0238]	[.0236]	[.0263]	[.0322]	[.0301]	[.0355]
Log Salary$_{t-2}$	.0101	-.0129	-.0018	***.3495	***.3593	***.2993	***.5603	***.5480	***.5288
	[.0184]	[.0218]	[.0224]	[.0302]	[.0302]	[.0537]	[.0845]	[.0734]	[.0778]
N	735	735	735	1522	1522	1520	988	988	988
Log Liklihood / F-stat	.5497	-474.54	-468.04	9.04	7.70	6.04	3.31	7.04	7.39
Pseudo R^2	.0669	.0628	.0665	.1042					

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test); † = significant at $p \leq 0.1$ (two-sided test)

Appendix.

Table A1. Prospective analysis of current salary on likelihood of moving into self-employment in subsequent periods for PhD subsample. The sample includes males whose responses are included in the SDR restricted file in 1995, 1997, 1999, and 2001 and who were at least 22 in 1995 and not more than 65 in 2001. Individuals in the SDR file who were not in the labor force in all 4 periods are eliminated from the sample. Individuals whose PhDs were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. The dependent variable is the log of the weekly wage at time t. All regressions are conditional on not being self employed at time *t-2*. All covariates are at time *t*, unless otherwise specified. For the OLS regression, robust standard errors are presented (in brackets). For the quantile regressions, bootstrap standard errors with 100 repetitions are presented. N = 13,758.

Analysis Series	Key Coefficients	OLS	Quantile: 10%	Quantile: 25%	Quantile: 50%	Quantile: 75%	Quantile: 90%
(1)	$SELF\text{-}EMPLOYED_{t+2}$	-.0366 [.0344]	**-.1667 [.0548]	-.0456 [.0326]	.0067 [.0233]	*.0691 [.0328]	*.1073 [.0442]
(2)	$SMALLFIRM_t * SELF\text{-}EMPLOYED_{t+2}$	†-.0842 [.0496]	*-.2137 [.0899]	*-.1076 [.0461]	-.0096 [.0357]	.0697 [.0652]	†.1129 [.0694]
	$LARGEFIRM_t * SELF\text{-}EMPLOYED_{t+2}$	.0491 [.0343]	-.1181 [.0731]	.0154 [.0271]	.0462 [.0318]	†.0691 [.0375]	†.1052 [.0677]
(3)	SE_INC_{t+2}	.0431 [.0426]	-.0942 [.0599]	.0168 [.0317]	.0029 [.0346]	**.1331 [.0478]	.1071 [.0658]
	$SE\_NOT\_INC_{t+2}$	*-.1364 [.0552]	†-.2696 [.1543]	*-.1726 [.0563]	-.0024 [.0387]	-.0231 [.0555]	*.1072 [.0517]
(4)	$SMALLFIRM_t * SE\_INC_{t+2}$	.0464 [.0605]	-.0215 [.1068]	.0147 [.0442]	.0207 [.0640]	**.2138 [.0747]	**.2126 [.0676]
	$SMALLFIRM_t * SE\_NOT\_INC_{t+2}$	***-.2578 [.0779]	**-.5361 [.1862]	***-.2748 [.0629]	-.1042 [.0677]	-.1067 [.0659]	.0436 [.0858]
	$LARGEFIRM_t * SE\_INC_{t+2}$	.0217 [.0376]	-.1045 [.1051]	.0178 [.0263]	.0155 [.0344]	.0458 [.0375]	.0004 [.0539]
	$LARGEFIRM_t * SE\_NOT\_INC_{t+2}$	.0780 [.0584]	-.1198 [.1017]	.0106 [.0668]	†.0939 [.0546]	†.1467 [.0817]	**.1941 [.0739]
(5)	$SMALLFIRM_t * RNDTRACK_t * SELF\text{-}EMPLOYED_{t+2}$	†-.1720 [.0960]	-.1767 [.1523]	-.0378 [.0612]	†-.0935 [.0503]	-.0490 [.0761]	-.0087 [.0784]
	$SMALLFIRM_t * NOTRNDTRACK_t * SELF\text{-}EMPLOYED_{t+2}$	-.0500 [.0568]	*-.2657 [.1327]	**-.1604 [.0593]	-.0071 [.0353]	.1361 [.0873]	**.2162 [.0783]
	$LARGEFIRM_t * RNDTRACK_t * SELF\text{-}EMPLOYED_{t+2}$	†.0588 [.0324]	.0605 [.0473]	.0462 [.0327]	.0352 [.0431]	.0511 [.0468]	.0006 [.0687]
	$LARGEFIRM_t * NOTRNDTRACK_t * SELF\text{-}EMPLOYED_{t+2}$	.0430 [.0527]	**-.2402 [.0773]	.0024 [.0668]	.0479 [.0564]	.0833 [.0602]	*.1817 [.0891]

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test); † = significant at $p \leq 0.1$ (two-sided test)

Notes: Variables in wage equation not displayed above include controls for the state in which the employee works, the employee's highest degree, the field of the highest degree, age and age squared, job tenure and job tenure squared, race dummy variables, small firm dummy variable, and year dummies.

Table A2. Prospective analysis of current salary on likelihood of moving into self-employment in subsequent periods for non-PhD subsample. The sample includes males whose responses are included in the SESTAT restricted file in 1995, 1997, 1999, who were at least 22 in 1995 and not more than 65 in 2001, and who did *not* have PhDs. Individuals in the SESTAT file who were not in the labor force in all 4 periods are eliminated from the sample. Individuals whose highest degrees were not in a science or engineering field are also eliminated from the sample, as are all individuals who reported working fewer than 30 hours per week on average and fewer than 30 weeks per year. The dependent variable is the log of the weekly wage at time t. All regressions are conditional on not being self employed at time *t*. All covariates are at time *t*, unless otherwise specified. For the OLS regression, robust standard errors are presented (in brackets). For the quantile regressions, bootstrap standard errors with 100 repetitions are presented. N = 23,849

Analysis Series	Key Coefficients	OLS	Quantile: 10%	Quantile: 25%	Quantile: 50%	Quantile: 75%	Quantile: 90%
(1)	$SELF\text{-}EMPLOYED_{t+2}$	-.0058 [.0248]	***-.1689 [.0462]	-.0271 [.0326]	**.0471 [.0179]	***.0779 [.0161]	***.1899 [.0352]
(2)	$SMALLFIRM_t$ *$SELF\text{-}EMPLOYED_{t+2}$	-.0201 [.0364]	***-.2511 [.0670]	-.0553 [.0414]	**.0732 [.0261]	***.1236 [.0321]	***.2114 [.0370]
	$LARGEFIRM_t$ *$SELF\text{-}EMPLOYED_{t+2}$	-.0157 [.0294]	-.0220 [.0608]	-.0075 [.0325]	-.0167 [.0239]	*.0434 [.0206]	†.1030 [.0591]
(3)	SE_INC_{t+2}	.0322 [.0322]	-.0484 [.0558]	-.0171 [.0388]	**.0692 [.0249]	***.1089 [.0251]	***.2111 [.0398]
	$SE\_NOT\_INC_{t+2}$	-.0633 [.0387]	***-.3183 [.0697]	**-.1441 [.0495]	.0089 [.0282]	†.0483 [.0270]	*.1174 [.0551]
(4)	$SMALLFIRM_t$ *SE_INC_{t+2}	.0369 [.0453]	*-.1620 [.0809]	.0267 [.0406]	*.1113 [.0356]	***.1872 [.0387]	***.2444 [.0345]
	$SMALLFIRM_t$*$SE\_NOT\_INC_{t+2}$	†-.1131 [.0586]	***-.3944 [.0807]	***-.1978 [.0593]	-.0223 [.0459]	.0180 [.0415]	.0739 [.0806]
	$LARGEFIRM_t$*SE_INC_{t+2}	.0214 [.0403]	.0242 [.0359]	-.0144 [.0371]	.0245 [.0303]	.0263 [.0267]	.0295 [.0728]
	$LARGEFIRM_t$*$SE\_NOT\_INC_{t+2}$	.0078 [.0425]	**-.2478 [.0967]	-.0696 [.0691]	.0015 [.0462]	.0570 [.0524]	*.1468 [.0885]
(5)	$SMALLFIRM_t$ *$RNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	-.0434 [.0808]	†-.2539 [.1473]	-.0587 [.1158]	.0225 [.0425]	.0534 [.0647]	.1161 [.0974]
	$SMALLFIRM_t$ *$NOTRNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	-.0151 [.0404]	***.2493 [.0715]	-.0511 [.0449]	*.0825 [.0337]	***.1337 [.0341]	***.2202 [.0398]
	$LARGEFIRM_t$ *$RNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	.0071 [.0384]	.0875 [.1081]	.0425 [.0531]	.0006 [.0325]	.0139 [.0418]	-.0216 [.0549]
	$LARGEFIRM_t$ *$NOTRNDTRACK_t$* $SELF\text{-}EMPLOYED_{t+2}$	.0172 [.0343]	-.0450 [.0704]	-.0163 [.0417]	.0259 [.0295]	*.0484 [.0242]	**.1908 [.0728]

*** = significant at $p \leq 0.001$; ** = significant at $p \leq 0.01$; * = significant at $p \leq 0.05$ (two-sided test); † = significant at $p \leq 0.1$ (two-sided test)

Notes: Variables in wage equation not displayed above include controls for the state in which the employee works, the employee's highest degree, the field of the highest degree, age and age squared, job tenure and job tenure squared, race dummy variables, small firm dummy variable, and year dummies.