ESSAYS ON ENTREPRENEURSHIP AND EDUCATION

by

CHRISTOPHER J. YOUDERIAN

B.S., University of Idaho, 2007

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2012

Abstract

The first essay tests whether the returns to education are different between entrepreneurs and regular employees. If the signaling model of education is correct, entrepreneurs should receive lower returns from education (relative to employees) because they have no need to signal their productivity to an employer. However, this result should only hold if the researcher is able to control for selection into self-employment and the endogeneity of education. This is illustrated using a stylized model of signaling. The relationship between self-employment and the returns to education is tested using data from the 1996 Survey of Income and Program Participation. This rich panel dataset makes it possible to control for many business-specific characteristics, like business equity, that have been previously unaccounted for in the literature. Ordinary least squares regressions find the correlation between education and earnings to be weaker for entrepreneurs. To control for selection, I utilize a Heckman selection model using spousal health insurance and housing equity as instruments. It shows that selection biases downward the correlation between education and income for entrepreneurs. Finally, a fixed effects model is employed to control for any time invariant unobserved heterogeneity. This approach indicates that education is as valuable, if not more valuable, to entrepreneurs as it is to employees. This does not support the signaling hypothesis. The finding is robust to different measures of entrepreneurial earnings.

The second essay explores whether unemployed workers make successful transitions into self-employment. It is well established that unemployed workers are more likely to transition into self-employment than individuals coming from paid employment. A growing body of literature suggests that these formerly unemployed entrants tend to exit self-employment earlier than typical entrants. It is tempting to attribute this result to differences in ability between the two groups. However, using an adapted version of Frank (1988)'s Intertemporal

Model of Industrial Exit, I show that this is not the case. In this model, entrants to self-employment receive noisy information about their true entrepreneurial ability from their earnings in the market. I show that low ability entrants to entrepreneurship should be no more likely to exit self-employment than high ability entrants to self-employment. This is because although low ability entrants will earn less as entrepreneurs, their outside wage in paid employment will also be proportionately lower. Survival in self-employment, therefore, is a function of how initial expectations match reality. This leads me to suggest that the high exit rates out of self-employment for the formerly unemployed may be because this group systematically overestimates their entrepreneurial ability at entry. This hypothesis is justified by evidence from the psychology literature that low ability individuals tend to overestimate their performance. Duration analysis on data from the 1996 and 2001 panels of the Survey of Income and Program Participation confirms that the formerly unemployed are more likely to exit self-employment. I also find preliminary evidence consistent with the hypothesis that the unemployed overestimate their likelihood of success in self-employment. These findings should give policymakers pause before incentivicing the unemployed to enter self-employment.

ESSAYS ON ENTREPRENEURSHIP AND EDUCATION

by

CHRISTOPHER J. YOUDERIAN

B.S., University of Idaho, 2007

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2012

Approved by:

Major Professor
Dong Li

Copyright

Abstract

The first essay tests whether the returns to education are different between entrepreneurs and regular employees. If the signaling model of education is correct, entrepreneurs should receive lower returns from education (relative to employees) because they have no need to signal their productivity to an employer. However, this result should only hold if the researcher is able to control for selection into self-employment and the endogeneity of education. This is illustrated using a stylized model of signaling. The relationship between self-employment and the returns to education is tested using data from the 1996 Survey of Income and Program Participation. This rich panel dataset makes it possible to control for many business-specific characteristics, like business equity, that have been previously unaccounted for in the literature. Ordinary least squares regressions find the correlation between education and earnings to be weaker for entrepreneurs. To control for selection, I utilize a Heckman selection model using spousal health insurance and housing equity as instruments. It shows that selection biases downward the correlation between education and income for entrepreneurs. Finally, a fixed effects model is employed to control for any time invariant unobserved heterogeneity. This approach indicates that education is as valuable, if not more valuable, to entrepreneurs as it is to employees. This does not support the signaling hypothesis. The finding is robust to different measures of entrepreneurial earnings.

The second essay explores whether unemployed workers make successful transitions into self-employment. It is well established that unemployed workers are more likely to transition into self-employment than individuals coming from paid employment. A growing body of literature suggests that these formerly unemployed entrants tend to exit self-employment earlier than typical entrants. It is tempting to attribute this result to differences in ability between the two groups. However, using an adapted version of Frank (1988)'s Intertemporal

Model of Industrial Exit, I show that this is not the case. In this model, entrants to self-employment receive noisy information about their true entrepreneurial ability from their earnings in the market. I show that low ability entrants to entrepreneurship should be no more likely to exit self-employment than high ability entrants to self-employment. This is because although low ability entrants will earn less as entrepreneurs, their outside wage in paid employment will also be proportionately lower. Survival in self-employment, therefore, is a function of how initial expectations match reality. This leads me to suggest that the high exit rates out of self-employment for the formerly unemployed may be because this group systematically overestimates their entrepreneurial ability at entry. This hypothesis is justified by evidence from the psychology literature that low ability individuals tend to overestimate their performance. Duration analysis on data from the 1996 and 2001 panels of the Survey of Income and Program Participation confirms that the formerly unemployed are more likely to exit self-employment. I also find preliminary evidence consistent with the hypothesis that the unemployed overestimate their likelihood of success in self-employment. These findings should give policymakers pause before incentivicing the unemployed to enter self-employment.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am grateful for all of the assistance I received on this dissertation and throughout the PhD program. In particular, I would like to thank my committee for their comments and guidance. To my chair, Dr. Dong Li, thank you for being patient, pushing me when necessary and always being supportive. I truly appreciate all of the time you have invested in me throughout this process. Your emphasis on making a real contribution to the literature and your ability to be both straightforward and kind are qualities I aspire to emulate. I am also indebted to Dr. Dennis Weisman, Dr. Philippe Belley, and Dr. Jeff Peterson for their thought-provoking comments, challenging questions, and encouragement.

I am also deeply appreciative of many other professors who helped me get to this point. Dr. Ashley Lyman ignited my interest in economics as an undergraduate and encouraged me to apply for graduate school. Dr. Lloyd B. Thomas took a chance on me as a late applicant to the program. Dr. Bill Blankenau offered helpful advice and always had the best interests of the graduate students at heart. Dr. Florence Neymotin taught me much about labor economics and applied research. I also am thankful to Dr. James F. Ragan, for his kindness when I was his graduate teaching assistant. His example as a researcher and teacher taught me much and I wish that we had more time with him.

I want to express my gratitude to my friends and colleagues who made graduate school so much more enjoyable. To Vladimir Bejan, with whom I shared an office and many late nights. To Mark Melichar, whose conversations and coffee made the day more enjoyable and whose work kept GSE from falling apart. The camaraderie and friendship of Jenny Wu, Yunyun Lv, Jessie Xie, Shu Zhou, Aditi Bothra, Abhinav Alakshendra, Patrick Scott, Bebonchu Atems, Ben Le, Mercy Palamuleni, Scott Unekis, Dave and Jamie Brown, Casey Abington, Yaseen Alhaj-Yaseen and many others made all the difference. I am also thankful for everyone who played on our economics intramural basketball teams including "Too Big to Fail." In particular, I am grateful to Dr. Daniel Kuester for acting as team strategist and keeping our sportsmanship rating playoff-eligible.

Finally, I am thankful to my family for their love and encouragement. To my wife and fellow economist Laura for her love, support, proofreading, and sense of humor. She gave me substantial feedback on this dissertation and kept me motivated when I was frustrated. I cannot imagine going through this process without her by my side. To my parents for loving me and fostering my curiosity. And, to Andrew and Carrie for the many ways in which they have shaped and encouraged me over the years.

Dedication

To Laura.

Chapter 1

The Returns to Education for Entrepreneurs

1.1 Introduction

Entrepreneurship receives a lot of attention. Politicians tout small businesses as the driving force of job creation. The media glamorize technological start-ups and their young founders. Hollywood profiles their lives. And, economists generally agree that entrepreneurs contribute to innovation, create jobs, and increase productivity (Van Praag and Versloot, 2008). Many economics textbooks even include entrepreneurship as a factor of production along with labor, capital and land. But, for all its prominence, entrepreneurship is poorly understood and, until recently, has been largely neglected in the economics literature (Parker, 2004). Labor market surveys often fail to collect detailed information about the self-employed. The data that is collected is often ignored by labor economists studying typical workers. This is because the self-employed only represent a fraction of the population and cannot be easily compared to employees. Their earnings and hours are not only highly volatile, but subject to measurement error.

In fact, there is even disagreement within the field as to who should be considered an entrepreneur. Entrepreneurship is often associated with bringing new innovations to market. Such creativity often originates from small businesses. Thus, entrepreneurship is loosely used to describe anyone who is self-employed. Critics argue that this is a mistake as it includes many small businesses like dentists and taxi drivers that are not especially innovative. As it is impossible to measure the "innovativeness" of a business, this paper adopts the more rudimentary definition of an entrepreneur as someone who is self-employed and uses the two terms interchangeably. The self-employed are of interest, even if they are not particularly innovative, because they have complete autonomy from an employer. This makes it possible to draw interesting inferences by comparing the self-employed to regular employees.

1.1.1 Motivation

This paper compares the returns to education for entrepreneurs to those for employees. The motivation for this topic is twofold. First, measuring the returns to education for entrepreneurs is important in its own right. There is serious public skepticism as to the value of education for entrepreneurs. In 2011, billionaire Paypal cofounder Peter Thiel founded a fellowship program that awarded twenty 100,000 dollar awards to 20 aspiring entrepreneurs under the age of 20. But the money came with a catch. These young award recipients were required to drop out of school for at least two years while they pursued their entrepreneurial venture.[1] The media often highlight anecdotal evidence to suggest that education is not as important for entrepreneurs. This narrative is aided by stories of college dropouts like Bill Gates, Steve Jobs, and Mark Zuckerberg who went on to start wildly successful businesses. This study will help determine whether education is as good of an investment for entrepreneurs as it is for employees. This is important not only for prospective entrepreneurs, but policymakers as well. Given its connection to innovation and job growth, governments routinely encourage self-employment. For example, the United States' Small Business Administration provides counseling and subsidized loans for small businesses and entrepreneurs. A better understanding of the returns to education for entrepreneurs will allow governments to more efficiently promote entrepreneurship through public policy.

Secondly, comparing the returns to education of entrepreneurs and employees may pro-

[1] More information about the fellowship is available at http://www.thielfoundation.org/.

vide broader insights into the value of education. There are two models in the literature that explain the returns to education. The human capital model developed by Becker (1975) and Mincer (1974) indicates that returns to education accrue from developing human capital that makes a worker more productive. The signaling model introduced by Spence (1973) indicates that education can be valuable even if it does not affect productivity as it may provide information to employers about the unobservable characteristics of employees. Since, education is more costly for people of lower ability, a higher level of education is a signal to employers that a person is intelligent, hardworking, perseverant, or intellectually curious. Employees should enjoy returns from both human capital and signaling. But, as first noted by Wolpin (1977), entrepreneurs should only reap the benefits from human capital as they do not need to signal to an employer. Therefore, if the signaling model is true, the returns to education for entrepreneurs should be lower than those for employees.

1.1.2 Current Literature

In order to compare employees to entrepreneurs, it is important to understand the characteristics and motivations of the self-employed. There is a large consensus in the literature that the self-employed earn less than employees of comparable quality. Entrepreneurs have both lower initial earnings and lower earnings growth resulting in a 35 percent median pay differential after 10 years (Hamilton, 2000). Many researchers attribute this to the non-pecuniary benefits of being your own boss. This is supported by many studies that find the self-employed are happier than employees even after controlling for job and personal characteristics (Parker, 2004). Therefore, people with a preference for autonomy may self-select into entrepreneurship. If choices about education and entrepreneurial entry are made jointly, it becomes quite difficult to compare the returns to education between the two groups.

A subset of the entrepreneurship literature has examined the returns to education for entrepreneurs. A meta-analysis of over 100 studies by Van der Sluis, Van Praag and Vijverberg (2008) finds that the returns to education for entrepreneurs are estimated to be

6.1 percent per year of schooling. This is somewhat lower than estimates for the returns to education in general. Of the 20 studies specifically examining the returns to education, the results were more mixed. U.S. studies tend to find that the returns to education are higher for entrepreneurs while European studies tend to find the returns to education are higher for employees. However, most of these studies are plagued with three common shortcomings: they fail to account for the endogeneity of education, ignore selection into entrepreneurship, and inappropriately measure the income of entrepreneurs. A few papers have attempted to credibly deal with some of these issues. Van der Sluis, Van Praag and Vijverberg (2009) use panel data from the 1979 NLSY[2] which provides them with controls on ability and family background instruments for education. Contrary to the signaling theory, they find that the returns to education for entrepreneurs are higher than for employees. Heywood and Wei (2004) study entrepreneurs in Hong Kong and correct for selection into entrepreneurship. They find the returns to education are significantly smaller for the self-employed. Brown and Sessions (1998) use Italian data and a bivariate selection model. They find evidence that the self-employed receive lower returns to education. This paper joins the former papers in an effort to credibly measure the returns to education for the self-employed by addressing these issues.

The signaling and human capital literature is extensive and there is still disagreement as to the role each plays in the returns to education. And yet, the question is of great importance. Economists typically consider education to convey positive externalities to society and thus have a social value that exceeds its private value. This implies that education is underproduced and should be subsidized. In contrast, if people acquire an inefficiently high level of education in order to signal their productivity, education is overproduced relative to what is socially optimal.[3] The debate is difficult to resolve because both theories yield the same empirical prediction: higher educated people should earn higher wages. Some

[2]The National Longitudinal Study of Youth follows people from 1979 annually and later biennially to present.

[3] Spence (2002) shows that in the presence of inefficient signaling, social welfare can be improved by taxing education.

researchers suggest that accurately controlling for ability is sufficient to show that the returns to education are driven by human capital accumulation. But, as long as firms do not know the ability of prospective workers, the signaling model predicts that the education coefficient will capture the inferences firms make from education about worker productivity (Weiss, 1995). Thus, economists have resorted to comparing screened and unscreened groups to evaluate the two schooling theories.

Wolpin (1977) spearheads this approach by predicting that if the signaling model is correct, education should be less valuable for people who are self-employed. Thus, he compares the education levels of paid workers and the self-employed. Contrary to the signaling theory, he finds that workers in the two groups obtained roughly the same amount of schooling. A slightly different approach is taken by Riley (1979). He notes that signaling should be much more important in sectors of the economy where worker productivity cannot be easily observed. In these screened sectors of the economy he predicts, and finds, that the lifetime earnings of the self-employed should be higher since they do not have to waste resources signaling their productivity. This paper explores the idea that if education is largely predetermined prior to the self-employment decision, the relevant question becomes whether the returns to education are higher or lower for the self-employed.

1.2 Theory

This paper tests for evidence of job market signaling by comparing the returns to education of entrepreneurs and employees.[4] The theoretical basis for this paper can be stated quite simply. Entrepreneurs are an unscreened[5] group because, since they are self-employed, they have no need to signal their ability to an employer. Employees, on the other hand, are

[4]I am indebted to Dr. Philippe Belley for challenging and helping me to think more deeply about the theoretical predictions in this paper. Any mistakes are my own.

[5]It is worth noting the differences between signaling and screening. Signaling is an effort by the party that is privy to private information (employees) to convey that information to an uninformed third party (employers). Screening is the use of this information by employers to make inferences about the unobservable characteristics of employees.

a screened group and must signal their productivity to employers. Therefore, education is valuable to employees because it both improves human capital and acts as a signal of unobservable characteristics. In contrast, education is only valuable to entrepreneurs as a source of human capital. Since both groups enjoy the human capital benefits of education, but only employees get the signaling benefits, the signaling model predicts that employees should receive higher returns from education than entrepreneurs. If the signaling model of education is incorrect, there should be no difference in the returns to education between the two groups. These theoretical predictions have been accepted in the literature as the logical extension of Wolpin (1977) to the returns to education. However, I show that the returns to education should only be lower for entrepreneurs if unobservable individual characteristics are held constant and signaling is imperfect.

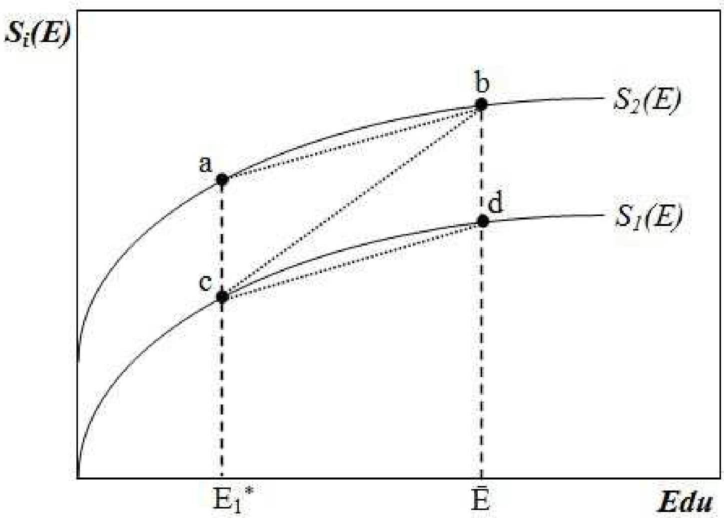
To illustrate these points, I utilize a stylized model. Many of the seminal papers on signaling (Spence, 1973, Riley, 1979) show that signaling can explain the returns to education even if education has no effect on human capital accumulation. However, this does not imply that signaling and human capital are mutually exclusive. Education can both augment human capital and provide value as a signal. Spence (2002) presents a parsimonious two-group model where education both enhances human capital and serves as a signal. This model, found in Section II of Spence (2002), is as follows: Let there be two types of workers, i = 1,2; E is the education chosen by each worker; $s_i(E)$ is the value of each worker; $s_i^{'}(E) > 0$ and $s_i^{''}(E) < 0$; $c_i(E)$ is the cost of education for each worker; $c_i^{'}(E) > 0$ and $c_i^{''}(E) > 0$; $s_2(E) > s_1(E)$; $c_1(E) > c_2(E)$.

Workers in group 1 are of a lower ability than workers in group 2 and have a higher cost of acquiring education. For both groups, $s_i(E)$ is a concave function and $c_i(E)$ is a convex function. Employers do not know to which group a potential employee belongs. However, since individuals in group 2 have lower costs of education, they will consume higher levels of education. Therefore, employers can use education to screen for ability. Workers in group 1 have an incentive to acquire extra education to imitate high ability group 2 workers.

However, since the costs of acquiring education are higher for group 1, in a separating equilibrium[6], group 2 workers will choose a level of education that is prohibitively costly for group 1 to obtain. The dynamics of this model can bee seen in Figure 3 of Spence (2002).

In this model, each worker maximizes the net return of education: $s_i(E)$ - $c_i(E)$. However, as researchers, we are largely unable to measure the costs of education. Instead, we only see the amount of education a worker has acquired and the wages they receive in the labor market. Figure 1.1 illustrates the relationship, implied by Spence (2002)'s model, that should exist between E and $s_i(E)$. The returns to education are captured by the slope of any line segment.

Figure 1.1: *Returns to Education Under Signaling*



E_1^* is the optimal amount of education that group 1 workers will choose so that $s_1(E)$ - $c_1(E)$ is maximized. $\overline{E}$ is the optimal amount of education that group 2 workers will choose such that it is not in the interest of group 1 workers to imitate them.[7] The slope of $\overline{cb}$ captures the returns to education for an employee who increases his education from E_1^* to

[6] A pooling equilibrium is possible if the fraction of the population in group 1 is sufficiently small. However, as this paper is testing for the presence of signaling, attention is restricted to the separating equilibrium.

[7] E_2^*, the optimal amount of education for group 2 in the absence of signaling, is less than $\overline{E}$ if signaling is inefficient.

$\overline{E}$. These returns are composed of two parts: 1) the employer now believes the worker is of higher ability and 2) the employer treats the increase in education as enhancing the human capital of a high ability worker. It should be noted that both group 1 and group 2 workers would get the same returns from this increase in education. However, it would only be a utility improving choice for group 2 because their education costs are sufficiently low. An entrepreneur increasing his education from E_1^* to $\overline{E}$ would enjoy returns to education equal to the slope of $\overline{ab}$ if he is a high ability worker and the slope of $\overline{cd}$ if he is a low ability worker. The returns to education for entrepreneurs are lower because they are not receiving a signaling benefit from education.

However, a researcher cannot typically identify the ability of individuals. The researcher simply sees workers' education choices. Therefore, while the model predicts the returns to education should be lower for entrepreneurs, this does not imply that the simple correlation between education and income should be smaller for the self-employed.

Reconsider the original model where signaling is perfect. The crucial point is that employees are paid their marginal product. The correlation between education and earnings for employees is the slope of $\overline{cb}$. This captures that highly educated employees have high innate ability which is recognized and rewarded by employers (signaling). Moreover, employers compensate workers for the impact education has on enhancing human capital.

Similarly, a researcher studying entrepreneurs is unable to discern the ability of individuals and therefore has the same perspective as an employer. Highly educated entrepreneurs have high innate ability and entrepreneurs with little education have low ability. Therefore, the correlation between education and earnings for entrepreneurs is also the slope of $\overline{cb}$. In other words, even though entrepreneurs do not receive the signaling benefits of education via an employer, from the researcher's perspective, educated entrepreneurs are also of higher ability and can be thought of as getting the "signaling" reward from the market. Therefore the correlation between education and earnings should be the same for entrepreneurs and employees. Intuitively, if the correlation between education and entrepreneurial earnings captures the true value of education, rational employers must reward education similarly.

This result is not conditional on perfect signaling. Suppose that education is an imperfect signal. Perhaps education is only 90 percent correlated with ability instead of 100 percent correlated with ability. Then, employees would receive less signaling value from their education. Similarly, researchers of entrepreneurs would see a weaker relationship between education and earnings as some high ability entrepreneurs would not get an education and do abnormally well.[8] In both cases, the correlation between education and earnings would fall equally. An example where education is an imperfect signal is illustrated in Figure 1.2 where the correlation between education and earnings is denoted by the slope of $\overline{ef}$.

In fact, the model's predictions only hold when signaling is imperfect. If signaling is perfect, all type 1 individuals will acquire E_1^* and all type 2 individuals will acquire $\overline{E}$. Both employees and entrepreneurs will be paid exactly their marginal product. If signaling perfectly captures differences in ability, after the researcher controls for ability, there will be no remaining signaling value of education. In contrast, when signaling is imperfect, employers will make inferences about ability from education that are correct on average for the population, but incorrect for those individuals who face barriers to acquiring education. If these individuals are employees, from the employer's perspective there is a very high probability they are of low ability and they will be paid less than their marginal product. However, if these individuals are entrepreneurs, they will be paid their true marginal product.

This raises the point that individuals who face barriers to acquiring education may have an incentive to transition into self-employment where the market will pay them their marginal product. To explore the impact of this type of selection more formally, consider the following exercise.

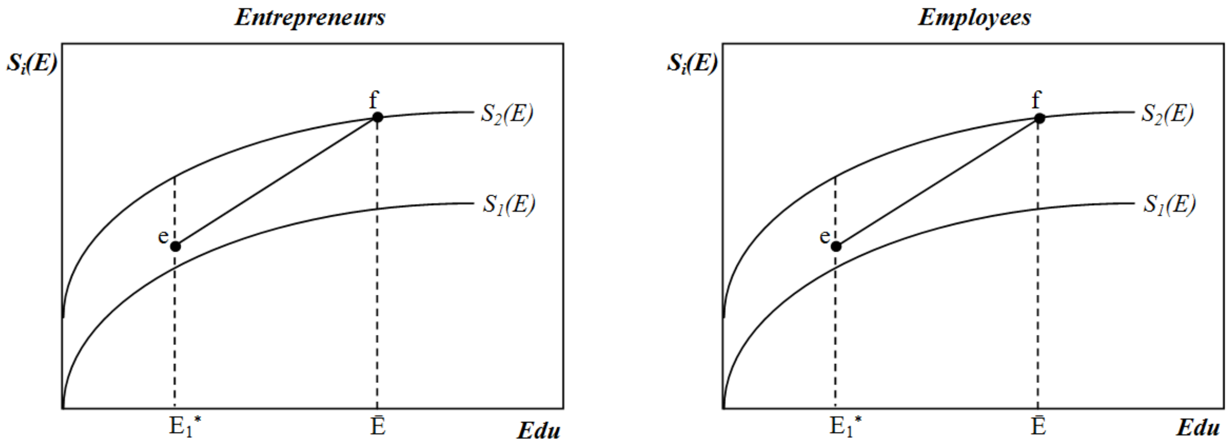
Suppose that there are 400 workers: 200 are self-employed and 200 are employees. In each group, 100 workers are high ability and 100 workers are low ability.[9] Suppose, that

[8]It is also possible that some low ability individuals would get so much education as to appear to be high ability. The effect on the returns to education would be identical. This seems less likely in equilibrium since the cost of acquiring education must be high enough to make it irrational for low ability copycats.

[9]For simplicity, the self-employed make up half of the population. The effects of selection will actually

there are barriers to signaling that prevent 25 percent of high ability workers from signaling via education. For example, maybe some high ability students face credit constraints on college financing. This will cause education to be a weaker signal of ability. Individuals with E_1^* are of low ability with probability .8 and high ability with probability .2.[10] Individuals with $\overline{E}$ are high ability with probability 1.0, but fewer individuals will now accumulate $\overline{E}$. The result is that low ability workers with little education will be paid more than their marginal product and high ability workers with little education will be paid less than their marginal product. This is illustrated in Figure 1.2 where e is above the marginal product of a low ability worker but below the marginal product of a high ability worker.

Figure 1.2: *Correlation between Education and Earnings under Imperfect Signaling*



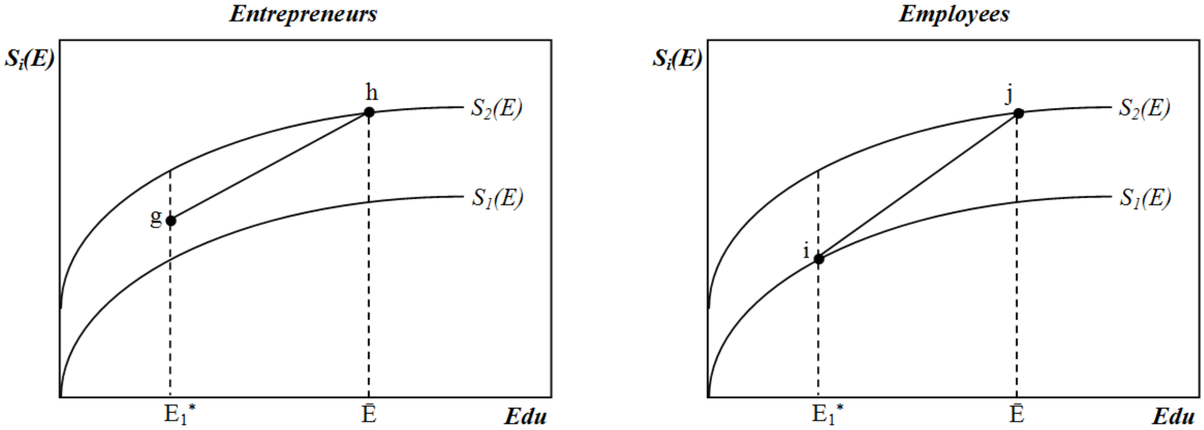
If self-employment allows a worker to receive his true marginal product, then high ability workers with little education will have a financial incentive to transition into self-employment. If all high ability, low education workers transition into self-employment the result is as illustrated in Figure 1.3. For employees, education becomes a perfect signal of ability. The returns to education is the slope of line segment $\overline{ij}$. However, self-employed individuals with E_1 are of low ability with probability 2/3 and high ability with probability 1/3.[11] Therefore, low education entrepreneurs have a higher average ability than low

be stronger if the self-employed make up a small fraction of the population as is the case empirically.

[10] 125 workers in each group will now acquire E_1^*. Of these, 25 will be high ability individuals.

[11] 150 employees will now acquire E_1. Of these, 50 will be high ability individuals.

Figure 1.3: *Imperfect Signaling and Selection*



education employees. This reduces the correlation between education and earnings for entrepreneurs. The correlation between education and earnings for entrepreneurs is the slope of $\overline{gh}$. From the researcher's perspective, the correlation between education and income should be much smaller for entrepreneurs than for employees. The extent to which this happens depends on the transferability of skills from paid to self-employment, the cost of switching to self-employment, and the knowledge and foresight of workers. An indication of this type of selection is that entrepreneurs should have lower levels of education.

In summary, in the presence of signaling entrepreneurs should earn lower returns from education only under certain conditions. First, the researcher must control for differences in ability caused by the endogenous nature of education. Otherwise, there is no reason the researcher should find the correlation between education and earnings to differ between the self-employed and employees. Second, barriers to signaling must exist which make education an accurate proxy of ability on average, but not for every individual. Finally, selection into self-employment by individuals unable to signal offers an alternative explanation for why entrepreneurs might see lower returns to their education. Therefore, in order to make inferences about signaling, it is necessary to control for possible unobservable heterogeneity caused by selection.

1.3 Methodology

1.3.1 Key Assumptions

The theoretical predictions depend on two additional assumptions: entrepreneurs do not use education as a signal and education is predetermined prior to the self-employment decision. Therefore, it is worth discussing these assumptions in greater detail.

The first assumption is based on the simple fact that the self-employed do not need to signal to an employer. However, education may be useful to the self-employed as a signal to creditors, investors, employees or customers. While this is a fair criticism, I argue that these alternative signaling avenues are relatively small. Small businesses are not financed only through bank loans. Instead entrepreneurs utilize credit cards, personal assets and loans from family and friends. In fact, many entrepreneurs report problems getting credit from a bank. This is likely caused by adverse selection in the loanable funds market. Banks have little information about the riskiness of new businesses. However, if they charge higher interest rates to compensate for this uncertainty only the riskiest businesses will be willing to borrow exacerbating the adverse selection problem. So, banks respond by rationing credit as described by Stiglitz and Weiss (1981). Assessing the risk of an individual loan is costly and banks are most likely to screen using metrics such as business income, credit history and collateral. In surveys, entrepreneurs report that a poor personal credit history and insufficient collateral are the most common reasons their loan applications are denied (Robb et al., 2008). Furthermore, an empirical study of U.S. small business owners in 1998 found no statistically significant relationship between education and loan approval (Vos et al., 2007). Investors may also screen on the basis of education. However, only 31.2 percent of the entrepreneurs in the SIPP are incorporated, a requirement for investment. And, investment is most common after a business has a track record, a much more informative signal than education. It is also possible that entrepreneurs must signal to customers in some industries. For example, customers may inquire about the education of an accountant or consultant. In addition, doctors and lawyers are required by law to attain a certain level of education in order to practice. However, an employee must signal both to his employer and to customers while an entrepreneur must only signal to customers. Finally, signaling to customers is the exception rather than the rule. In most industries consumers purchase a good or service with little knowledge or concern as to the education of the business owner. While education may hold some signaling value for entrepreneurs, I find the evidence to suggest that signaling is much more important for employees.

However, it is possible that education may play a unique role for entrepreneurs. Aldrich and Zimmer (1986) argue that entrepreneurs are not isolated actors and that social networks play a key role in their success. This seminal paper paved the way for a number of studies that examined the effect of entrepreneurial networks on firm success. Social networks have been credited with providing entrepreneurs market information, business advice, ideas, and emotional support (Hoang and Antoncic, 2003). Of particular interest, entrepreneurs seek legitimacy from social networks to reduce the implied risk to potential investors and employees. Cooper, Folta and Woo (1991) find some evidence that entrepreneurs with more education are better able to increase their social networks. It is, therefore, possible that education is valuable to entrepreneurs via social networks in ways that it is not valuable to employees. However, this effect is indirect and difficult to quantify.

The second assumption is that education is predetermined prior to the self-employment decision. If, on the contrary, children anticipate being entrepreneurs as adults they may intentionally underinvest in education. This was the original prediction made by Wolpin (1977). However, he finds the opposite result: entrepreneurs actually accumulate more education than employees. This finding is now well established in the literature. Similarly, the very fact that people switch from regular employment to self-employment and back again is inconsistent with a model where education choices are conditional on future self-employment. Decisions to invest in high school and college education are made when an individual is still an adolescent. It is unlikely, at this young age, that a student can predict with any certainty whether he or she will be self-employed later in life. In fact, Blanchflower and Oswald (1998) find virtually no correlation between childhood psychological characteristics and entry into self-employment as an adult.[12] Moreover, a meta analysis of 20 studies done by Van der Sluis, Van Praag and Vijverberg (2008) finds that education level has an insignificant effect on entry in 75 percent of the studies. The evidence suggests that future entrepreneurs accumulate about the same level of education as future employees. It may be that aspiring entrepreneurs either believe education is equally beneficial for the self-employed or they are acquiring education for consumption purposes or as a hedge against the risk of failing as an entrepreneur. The following sections explore selection and other methodological issues in more detail.

1.3.2 Methodological Issues

Selection

The first difficulty in comparing the returns to education between entrepreneurs and employees is selection into self-employment. Selection exists if entrepreneurs are systematically different from the type of people who choose to be employees. If people who are smarter, more independent, more creative, more risk loving, or more ambitious are more likely to become entrepreneurs, then the returns to education for entrepreneurs is capturing the interactions between education and these unobserved variables. The selection problem is reduced if personal characteristics like ability are observed in the data. But, unobserved characteristics like creativity are impossible to control for. To date, the literature finds little evidence of selection into entrepreneurship based on personal characteristics. Rather, there is strong evidence that capital constraints are the primary barrier to entry into self-employment. This is supported by evidence that individuals receiving an inheritance are more likely to enter self-employment (Blanchflower and Oswald, 1998). The same authors also find that capital constraints are the most common reason British survey respondents report for not becoming self-employed. Similarly, Evans and Jovanovic (1989) find a strong

[12] The only correlation they find is a weak negative relationship between being anxious for acceptance as a child and adult self-employment.

relationship between family assets and entry into self-employment in NLSY data. Later, I will utilize the impact of capital constraints on entry into self-employment to control for selection issues.

Endogeneity of Education

Another complication when comparing the returns to education between the two groups is the endogenous nature of the education. Education is endogenous because people with higher ability are more likely to acquire schooling. Therefore, the correlation between education and earnings is partially driven by unobservable differences between people who get more schooling and those who do not. As described in the theory section, the predictions of the theoretical model can only be tested if the researcher can control for differences in ability due to education. Card (1999) summarizes the various approaches that have been employed to estimate the causal effect of education on earnings. Most frequently, researchers measure the causal value of education by utilizing instrumental variables that arise from institutional rules or natural variation. For example, changes in state compulsory schooling laws and a student's quarter of birth (which influences the amount of schooling required before dropping out is allowed) have both been used as instruments. Similar instruments include the distance to a college and unemployment at the time of graduation. This group of instruments is often criticized for only being weakly correlated with education. Family characteristics like parents' education and number of siblings are more strongly correlated with education. But, these potential instruments are problematic as they tend to be correlated with earnings.

In this chapter, it is not ideal to correct for the endogeneity of education using instrumental variables. Finding strong instruments for education that are not correlated with income is quite difficult. Moreover, an instrumental approach provides the causal impact of education on those individuals who are compelled by the instrument to change their accumulation of education. It is quite possible that the instrument might have a different effect on the educational choices of employees than on the self-employed. Furthermore, the subset of individuals who are both self-employed and compelled to change their education based on the instrument is likely too small to generate meaningful estimates.[13] Instead, this paper employs a fixed effects model to control for individual heterogeneity when comparing the returns to education for employees and the self-employed.

1.3.3 Methodological Approach

Ordinary Least Squares

While my dataset, the SIPP[14], is a panel dataset, I begin by pooling the data and running an ordinary least squares regression. The coefficients will be biased if either education or selection into self-employment are endogenous. However, an OLS regression will provide insights beyond what is obvious from the summary statistics and act as a baseline that can be compared against the fixed effects results. The earnings equation, based on Mincer (1974), is as follows:

$$y_{it} = \beta_0 + Edu_i^{'}\beta_1 + (Edu_i * Self_{it})^{'}\beta_2 + \beta_3 Self_{it} + Controls_{it}^{'}\beta_4 + \epsilon_{it} \tag{1.1}$$

where y_{it} is log hourly wage of individual i at time t; Edu_i includes dummies for graduating high school, attending some college, graduating college, and completing a graduate degree; $Self_{it}$ is a dummy for being self-employed, $Edu_i * Self_{it}$ is an interaction of self-employment status and education dummies; $Controls_{it}$ include age, age^2, business and job tenure, percentage of time in the panel unemployed, regional dummies, incorporation of business status, reasons given for switching between jobs, and the whether a person is switching into or out of self-employment.

Heckman Selection Model

While OLS provides useful insights into the data, it only captures the correlation between education and earnings. This relationship will be biased if individuals who select into self-

[13] This statement is based on my fruitless modeling of both education and self-employment as endogenous variables where state unemployment at age 18 was used to instrument for education.

[14] Survey of Income and Program Participation.

employment are systematically different from those who chose paid employment. Almost by definition, entrepreneurs must be unique. Some unobserved characteristics must exist which allow entrepreneurs to see needs in the market and which compel them to fills those needs. To account for this, I employ a Heckman model to control for selection into entrepreneurship (Heckman, 1979). Often, the Heckman model is used to account for potential workers who decide not to work and therefore are not seen in the labor market data. In this case, both employee and entrepreneur labor market data is available. But, most individuals are only observed in one state: as either an entrepreneur or an employee. And, there are reasons to treat self-employment wages and paid employment wages as distinct dependent variables. It is quite possible that the factors that influence success in self-employment are quite different from those that influence success in the traditional labor market. For example, racial discrimination might have less of an effect on the self-employed if customers have less of a taste for discrimination than employers. Moreover, some variables that affect self-employment profit are not relevant for paid employees. Incorporation status and business tenure are not relevant to paid workers. Therefore, it makes sense to run separate regressions for entrepreneurs and employees and control for selection into the respective groups. Selection is controlled by using a probit model to determine what factors influence entry into self or paid employment. Identification of the Heckman model will be based on distributional assumptions unless additional variables are included in the first stage, but not the second stage. I employ two of these exclusion restrictions. The first, spousal insurance, captures whether an individual obtains health insurance from a spouse or other family member. It is well established that adverse selection in health insurance drives up the cost of coverage for people trying to acquire independent policies. Therefore, it is not surprising that less than 20 percent of small business owners get health insurance from their business.[15] This makes entrepreneurship more attractive to individuals who are able to obtain health insurance from their spouse (and consequently, paid employment is less attractive). The

[15] http://www.gallup.com/poll/101689/smallbusiness-owners-see-need-overhaul-us-healthcare.aspx

second instrument, home equity, captures the amount of equity the family has built up in their home. As explained earlier, there is strong evidence that capital constraints prevent many workers from starting their own business. Therefore, all else equal, individuals with more equity in their houses should be more likely to undertake an entrepreneurial venture. The first stage probit estimation for each model is given by:

$$\begin{aligned} Prob(Self = 1|Z) &= \Phi(Z'\gamma^{self}) \\ Prob(Employee = 1|Z) &= \Phi(Z'\gamma^{emp}) \end{aligned} \tag{1.2}$$

Where Z includes the original regressors, spousal insurance and log housing equity and Φ is the normal cumulative distribution function.[16]

These first stage results are used to construct the inverse mills ratio $\lambda_{it} = \phi(Z'\gamma)/\Phi(Z'\gamma)$ where ϕ is the population density function. The inverse mills ratio is then plugged into the regression to control for selection so the final estimated equation for each group is:

$$\begin{aligned} y_{it}^{self} &= \beta_0 + Edu_i^{'}\beta_1 + \beta_2\lambda_{it}^{self} + Controls_{self,it}^{'}\beta_3 + \epsilon_{it} \\ y_{it}^{employee} &= \alpha_0 + Edu_i^{'}\alpha_1 + \alpha_2\lambda_{it}^{employee} + Controls_{emp,it}^{'}\alpha_3 + \epsilon_{it} \end{aligned} \tag{1.3}$$

where $Controls$ exclude education dummies, $Controls_{self}$ contains entrepreneur specific variables like incorporation status, y_{it}^{self} is log hourly wage for the self-employed and y_{it}^{emp} is the log hourly wage for paid employees.

Fixed Effects Model

While the Heckman model is an improvement over OLS, it does not control for the endogeneity of education. As described in the theory section, when education is endogenous, there is no reason to expect the correlation between education and earnings to be lower for entrepreneurs. Therefore, I follow Van der Sluis, Van Praag and Vijverberg (2009) and employ the fixed effects model. A fixed effects model eliminates both the selection and endogeneity problem by holding unobservable individual effects constant. In the standard returns to

[16] State specific variables like business or job tenure and incorporation status are excluded because they exist only when state = 1.

education literature, fixed effects models are not used because education for typical workers is determined prior to entering the workforce. Because education is time invariant, it drops out during estimation leaving the returns to education unidentified. However, by interacting a dummy for self-employment with education, a fixed effects model is able to measure the premium (or penalty) on the returns to education for entrepreneurs. This premium (or penalty) is identified in the model by individuals who move from self-employment to regular employment and vice versa. The model is as follows:

$$y_{it} = \beta_0 + (Edu_i * Self_{it})'\beta_1 + \beta_2 Self_{it} + X_{it}'\beta_3 + Z_i + \mu_i + v_{it} \tag{1.4}$$

where y_{it}, $(Edu_i * Self_{it})$ and $Self_{it}$ are as identified in the ordinary least squares regression. Vector X captures time variant variables like age and marital status. Vector Z captures time invariant variables like race, gender and education that drop out during the within transformation.[17] μ_i captures the individual effect and v_{it} is the residual.

All time invariant variables including education drop out in the fixed effects model. But, I can identify β_1 which captures the premium (or penalty) the self-employed receive on the returns to their education. If the signaling model is correct, $\beta_1 < 0$ and education is less valuable for entrepreneurs. This coefficient is only identified through individuals who move into or out of self-employment in the sample.

In summary, the OLS model is used as a baseline. The Heckman model corrects for selection into entrepreneurship, but not for the endogeneity of education. It provides insights into the effects of selection on the returns to education as well as makes it possible to compare this paper with other papers that have employed the Heckman technique. Finally, the fixed effects model controls for all types of unobservable heterogeneity by holding individual characteristics constant and measuring the value of education as workers transition in and out of self-employment.

[17] Education is typically predetermined prior to entry into the labor market. To avoid measurement error issues, I replace every value of education for an individual with the highest level they ever report.

1.4 Data

This paper uses data from the Survey of Income and Program Participation 1996 (SIPP96). This dataset captures the movement of workers in and out of self-employment and has been used previously in the entrepreneurship literature. It contains a number of unique features that make it ideal to explore the question of interest.

The Survey of Income and Program Participation was created in the 1970's to supplement the CPS by providing a more accurate measure of income, employment, assets, health insurance and participation in government programs. While the CPS required respondents to report income on an annual basis, the SIPP surveyed households every 4 months. In 1996, the SIPP was redesigned to cover 40,000 households interviewed over the course of 4 years. This resulted in a sample of over 3 million monthly observations of working-age individuals. The self-employed are only a small subset of the workforce and only a fraction of them move into or out of self-employment in a given time frame. Therefore, a large dataset like the SIPP is required to generate sufficient variation to study the movement of workers into and out of self-employment.

Another benefit of the SIPP is that it contains very detailed information about the self-employed. Unlike other surveys, the SIPP asks the self-employed a number of specific questions about their business or businesses. Entrepreneurs are asked about the income they draw from up to two distinct business. Data are collected on the incorporation status of the business, the presence of a partner, historical earnings, expected earnings, the industry of the business, and the age of the business. If the worker leaves self-employment, the survey inquires as to the reasons for his/her departure. Every year, a topical survey records the equity an individual holds in his/her business. This is important since a large component of compensation for the self-employed is through the increased value of their business. Moreover, entrepreneurs may limit the salary they draw to grow their business or to avoid certain taxes.

While SIPP interviews are conducted three times a year, respondents are asked to recall information for each of the last 4 months. This format minimizes measurement error while providing a detailed employment history. Monthly data is useful to pinpoint exactly when someone moves into employment. It also makes it possible to determine whether someone moved directly from a job to self-employment (or vice versa) or whether a period of unemployment occurred between the switch. This is useful to control for selection issues.

Although the SIPP is ideal in many respects, it does have some limitations. For example, the SIPP contains monthly data but it only lasts for a period of 4 years. In such a short period of time only a small number of workers switch between regular jobs and self-employment. Moreover, the variation in any single worker's wage will undoubtedly be relatively small over the course of 4 years. The SIPP also provides little background information about respondents. It does not include a measure of ability or any information on family background. This makes pooled regressions of SIPP data subject to omitted variable bias. Finally, the SIPP contains categorical variables for degree completed instead of a continuous measure of education. This is the result of a decision in the 1980's by the U.S. Census Bureau to use degree-completed to measure post secondary education levels. Given the aims of this paper, using degree categories rather than years of education is preferred as degrees are more likely to act as a signal. Nonetheless, information on years of education would be desirable to run robustness checks.

1.5 Summary Statistics

Prior to comparing the returns to education between entrepreneurs and employees it is instructive to examine the differences between the two groups. Summary statistics for the self-employed and private paid workers can be found in Table 1.1. The unemployed, government workers, and non-profit workers are excluded from the sample. Similarly, part-time workers are dropped from the sample. These adjustments are made in an effort to compare two groups of full time workers who differ only in that some work for themselves and others work for a private employer.[18]

Table 1.1 shows that the self-employed are older, slightly better educated, more likely to be male, less likely to be a minority and more likely to be married. The self-employed work about 5.7 more hours per week on average. For their work, they earn \$20.66 per hour compared to the \$15.08 per hour earned by private employees. This gap widens considerably if the self-employed are credited with the increase in their business' equity over the month. While the self-employed earn more than their employee counterparts, these summary statistics alone cannot predict whether their premium in pay will remain after controlling for education and demographic differences between the two groups. Interestingly, the self-employed do not appear to be under-investing in education as Wolpin (1977) suggested they might if the returns to education are lower for the self-employed. Moreover, this is not consistent with the theoretical possibility of high ability individuals who are unable to acquire education flocking to self-employment.

Out of the workforce, 63.96 percent classify themselves as working for private employers and 13.56 percent classify themselves as self-employed.[20] This is a slightly higher frequency of self-employed workers than what others studies have found, but it is not unreasonably high. In a fixed effects model, the returns to education for entrepreneurs are identified only by those people who move into or out of self-employment (henceforth called switchers). Individuals who switch between the two categories at least once in the panel make up 15.0 percent of the workforce. It is natural to ask whether people who switch between self-employment and paid employment are systematically different from people who do not switch jobs. If there are serious differences between the two groups, selection may bias the results of a fixed effects estimation. Table 1.2 provides statistics on the differences be-

[18]Interestingly, the pay gap widens considerably when including part time workers. This may be because part time entrepreneurs choose only the most lucrative projects. Or, they may consider their work to be a side job and neglect to properly account for expenses. Regardless, for the purposes of this paper it makes sense to ignore this atypical group.

[20]Government and not-for-profit employees are excluded from analysis.

Table 1.1: *Self-employed vs. Employees*

	Self-employed	Employee	Difference	T-Statistic
Age (years)	45.39	41.28	4.11	109.60
Education (years)	13.67	13.46	0.21	20.20
High School (%)	0.30	0.31	-0.01	-4.25
Some College (%)	0.27	0.31	-0.04	-22.93
College Degree (%)	0.19	0.18	0.01	6.11
Graduate Degree (%)	0.13	0.10	0.03	27.93
Weekly Hours (%)	51.86	46.11	5.75	118.00
Male (%)	0.74	0.53	0.21	132.30
Black (%)	0.05	0.11	-0.06	-79.45
Hispanic (%)	0.07	0.10	-0.02	-26.18
Other (%)	0.05	0.04	0.00	6.07
Married (%)	0.75	0.64	0.11	72.78
Tenure (years)	10.47	7.39	3.07	91.06
Hourly Wage ($)	20.66	15.08	5.59	54.51
Hourly Wage with Equity ($)	42.27	15.08	27.19	76.82

Restricted to individuals over the age of 24 who are not in school or working part time.

Table 1.2: *Self-employed: Switchers vs. Non-switchers*

	Switchers	Non-Switchers	Difference	T-Statistic
Age	41.82	47.73	-5.91	-79.63
Education (years)	13.41	13.61	-0.20	-9.94
High School (%)	0.29	0.30	-0.02	-5.35
Some College (%)	0.31	0.27	0.04	13.63
College Degree (%)	0.18	0.19	-0.01	-4.32
Graduate (%)	0.10	0.13	-0.02	-11.03
Weekly Hours (%)	42.10	44.14	-2.04	-15.48
Male (%)	0.64	0.68	-0.03	-9.63
Black (%)	0.06	0.05	0.01	6.82
Hispanic (%)	0.10	0.06	0.04	20.04
Other (%)	0.04	0.04	0.00	-3.53
Married (%)	0.70	0.75	-0.06	-18.01
Hourly Wage ($)	22.21	24.98	-2.77	-4.27
Hourly Wage with Equity ($)	48.43	61.94	-13.51	-6.99
Tenure (years)	5.89	11.81	-5.92	-101.47
Incorporated Business (%)	0.22	0.27	-0.05	-17.37

These are self-employed individuals who at some time during the panel move from self-employment to regular employment or vice versa.

Table 1.3: *Employees: Switchers vs. Non-switchers*

	Switchers	Non-Switchers	Difference	T-Statistic
Age	41.09	41.00	0.09	1.30
Education (years)	13.36	13.05	0.31	16.58
High School (%)	0.28	0.34	-0.06	-20.90
Some College (%)	0.30	0.32	-0.01	-4.32
College Degree (%)	0.19	0.16	0.03	10.55
Graduate (%)	0.10	0.06	0.04	20.75
Weekly Hours (%)	41.50	42.32	-0.82	-8.10
Male (%)	0.63	0.53	0.09	29.22
Black (%)	0.07	0.10	-0.03	-18.36
Hispanic (%)	0.11	0.10	0.00	2.03
Other (%)	0.04	0.04	0.00	-1.81
Married (%)	0.66	0.63	0.02	8.07
Hourly Wage ($)	16.41	14.87	1.54	6.60
Tenure (years)	3.85	6.56	-2.71	-61.97

These are current employees who at some time during the panel move from self-employment to regular employment or vice versa.

tween self-employed switchers and self-employed non-switchers. On average, self-employed switchers are 5.9 years younger than and earn slightly less than those who are self-employed throughout the sample. Educational attainment is slightly lower for switchers but the differences are small. Table 1.3 shows the differences between employee switchers and employee non-switchers. For employees, the impact of switching is reversed. Switchers are about the same age as non-switchers, but they earn 1.54 dollars more per hour and have slightly higher levels of educational attainment. Switching during the panel is correlated with higher wages for employees and lower wages for the self-employed.

1.6 Results

1.6.1 Ordinary Least Squares

The results from the ordinary least square regression are available in Table 1.4. The results are best summarized by breaking the variables into categories.

Table 1.4: *Ordinary Least Squares*

Log hourly wage	(a) Base	(b) Equity	(c) No Prof.	(d) No Prof. + Equity
High School	0.178***	0.176***	0.179***	0.176***
	(0.004)	(0.004)	(0.004)	(0.004)
Some College	0.323***	0.321***	0.313***	0.312***
	(0.004)	(0.004)	(0.004)	(0.004)
College Degree	0.618***	0.618***	0.607***	0.605***
	(0.004)	(0.005)	(0.004)	(0.005)
Graduate Degree	0.843***	0.840***	0.798***	0.795***
	(0.006)	(0.006)	(0.006)	(0.007)
High School x Self	−0.048***	0.083***	−0.049***	0.082***
	(0.009)	(0.010)	(0.009)	(0.010)
Some College x Self	−0.036***	0.094***	−0.030***	0.099***
	(0.009)	(0.010)	(0.009)	(0.010)
College x Self	−0.102***	0.014	−0.108***	0.009
	(0.010)	(0.011)	(0.010)	(0.011)
Graduate x Self	0.071***	0.186***	−0.179***	−0.052***
	(0.011)	(0.013)	(0.013)	(0.015)
Self-Employed	−0.233***	−0.016	−0.227***	−0.01
	(0.009)	(0.010)	(0.008)	(0.009)
Incorporated	0.246***	0.309***	0.248***	0.314***
	(0.005)	(0.006)	(0.005)	(0.006)
Job Tenure	0.016***	0.016***	0.016***	0.016***
	(0.000)	(0.000)	(0.000)	(0.000)
Business Tenure	0.005***	0.011***	0.005***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)
Left Bus. Voluntarily	0.032***	0.079***	0.030***	0.081***
	(0.009)	(0.010)	(0.009)	(0.010)
Left Bus. Involuntarily	−0.101***	−0.102***	−0.094***	−0.099***
	(0.014)	(0.015)	(0.014)	(0.015)
Left Job Voluntarily	−0.040***	−0.016***	−0.036***	−0.012**
	(0.004)	(0.004)	(0.004)	(0.004)
Left Job. Involuntarily	−0.130***	−0.119***	−0.128***	−0.119***
	(0.004)	(0.004)	(0.004)	(0.004)
Switcher: Self to Emp.	−0.023*	−0.120***	−0.028**	−0.121***
	(0.009)	(0.010)	(0.010)	(0.010)
Switcher: Emp. to Self	0.079***	0.149***	0.079***	0.151***
	(0.010)	(0.010)	(0.010)	(0.010)
Total Switches	−0.098***	−0.098***	−0.101***	−0.110***
	(0.011)	(0.012)	(0.011)	(0.013)
(%) Unemployed	−0.068***	0.009	−0.054***	0.019**
	(0.005)	(0.006)	(0.005)	(0.006)
N. of Observations	520,626	506,598	502,262	489,210
R^2	0.207	0.201	0.2	0.188

Age, race, gender, and region variables omitted for space.

Demographics Age, male, white, and marriage are correlated with higher earnings. People who live in urban areas or in the Northeast or Western regions of the United States also receive a wage premium. These variables are omitted from the table due to space constraints.

Self-employment Penalty Consistent with the literature, being self-employed is accompanied with a 23.3 percent wage penalty.

Education High school graduates earn 17.8 percent more than dropouts. A college degree raises that premium to 61.8 percent. Those who receive graduate degrees earn 84.3 percent more than high school dropouts. However, the self-employed earn lower returns on their education. A high school diploma is worth an estimated 4.8 percent less to someone who is self-employed. Entrepreneurs who attend college find that their degree is worth 10.2 percent less than their employee counterparts. Interestingly, the relationship is reversed for the most highly educated individuals. A graduate degree is worth 7.1 percent more to an entrepreneur than to a paid worker. These results are curious since the theoretical model predicts that the correlation between education and income should only be lower for entrepreneurs after controlling for the endogeneity of education. Therefore, these results may be driven by selection into entrepreneurship as opposed to signaling.

Business Variables One of the advantages of the SIPP is that it contains data on a person's tenure at a job or with their business. Each year someone stays at a job increases his/her wage by 1.6 percent. This is consistent with the theory of firm-specific human capital. Similarly, self-employed wages increase by .56 percent for every year they stay with the business. Individuals that own an incorporated business enjoy a 24.6 percent wage premium over non-incorporated entrepreneurs. This is likely due to selection effects. More successful businesses are more likely to go through the hassle of becoming incorporated. However, incorporated entrepreneurs are able to reduce their immediate tax burden by only drawing a portion of their earnings as a salary and avoiding paying payroll taxes on earnings retained by the company. Therefore, it appears that the selection effects of incorporation dominate the tax incentive effects.

Panel Variables The regression also includes variables on why an individual left his/her last job or business. Leaving a business voluntarily is correlated with a 3.2 percent increase in wages while leaving a business involuntarily is correlated with a 10.1 percent decline in wages. Leaving a previous job is always associated with a wage penalty, but the penalty is substantially more severe if the job loss was involuntary.

Similarly, the regression controls for whether someone switched between self and paid employment during the panel and the direction of the switch. Consistent with the summary statistics, being a switcher is correlated with nearly a 10 percent decline in wages. This penalty largely disappears for people who are moving from employment into self-employment but is slightly exacerbated by individuals moving out of self-employment. This is consistent with individuals who fail at entrepreneurship being penalized when they return to the traditional labor force. Or, alternatively, it suggests that poor labor market prospects drive low ability workers into self-employment.

Alternative Specification: Added Equity

As discussed earlier, an entrepreneur's salary underestimates the marginal product of his labor if his productivity increases the value of the business. Therefore, even measures of profits will underestimate the value created by an entrepreneur if his firm is growing. The SIPP contains data on business equity yearly. I use linear imputation to create an estimate of the equity added each month. This is added to an entrepreneur's salary draw and divided by hours to determine hourly wage. This is, unfortunately, an imperfect solution. The value of a private business is based on the subjective opinion of the entrepreneur introducing the potential for substantial measurement error. Also, changes in business equity capture both the returns to an owner's labor and capital invested in the business. Finally, business equity is a household variable which may be double counted if two spouses own a business jointly. For these reasons, it is more appropriate to consider the hourly wage including added equity as the upper bound on the true wage a self-employed person is earning.

Given the above qualifications, incorporating added equity into log hourly wage has some significant consequences. It reduces the penalty on self-employment to less than 2 percent. It doubles the returns to business tenure and increases the coefficient on incorporation from 24.6 to 30.9 percent. Finally, it makes the interaction between self-employment and education positive for all degree categories (although college is not statistically significant).

Alternative Specification: Excluding Professionals

Part of the motivation for comparing the returns to education of entrepreneurs and the self-employed is that they represent screened and unscreened groups. Employers screen employees on the basis of education. The self-employed do not face such screening. However, as Wolpin (1977) and others in the literature have noted, some entrepreneurs may use education to signal to their customers. For example, professionals like doctors and lawyers often own their own practices and are thus self-employed. However, they are not able to legally practice unless they have attained a certain level of education. Consequently, their customers may very well base their choice of a doctor or lawyer on his or her educational credentials. Therefore, it would not be appropriate to consider self-employed doctors and lawyers to be an unscreened group.

Another problem with such professionals is that they often begin their careers working as an employee in a practice. Later, only after gaining experience and building a reputation, they take over the practice of a retiring practitioner or establish their own practice.[21] For this group, selection into self-employment is a proxy for career advancement and the interaction of self-employment and education is biased upward.

To avoid these issues I run a regression specification that excludes professionals in the fields of law and medicine. The results can be seen in Table 1.4. In the original regression,

[21] For example, when lawyers make partner in the firm they become self-employed and their salary jumps.

self-employed individuals with an advanced degree receive a 7.1 percent premium on their education. Excluding professionals turns that into a 17.9 percent penalty. It is problematic to subjectively pick and choose occupations where educational signaling may be important for the self-employed. Given that this issue will predominately affect workers with advanced degrees, the returns to education for this group should be interpreted with caution.

1.6.2 Heckman Selection Model

The first stage probit results can be seen in Table 1.5. Education and male are strongly correlated with selection into self-employment. Minorities are significantly less likely to become self-employed. Older workers are more likely to enter the workforce in general. Living in an urban area is correlated with an increased probability of being a paid employee. λ is statistically significant at the 95 percent level indicating that selection is indeed a problem. The first selection instrument, spousal insurance, is strongly positive for the self-employed and strongly negative for employees. The second instrument, log of housing equity, is also positive for the self-employed and negative for employees. A doubling of home equity makes a person about 1 percent more likely to be self-employed and 1 percent less likely to be an employee. This is consistent with my predictions about the relationship between health insurance, housing equity and employment choices.

The results from these second stage regressions can be seen in Table 1.6. Controlling for selection and running the regressions separately indicate that the returns to education are as high, if not a little higher, for entrepreneurs. This contradicts the earlier pooled OLS findings. A high school degree increases wages by 19.2 percent for the self-employed, but only 11 percent for a paid employee. A college degree is worth roughly the same to both groups. These findings suggest that unobservable differences between the self-employed are driving the low returns to education seen by entrepreneurs under OLS. One plausible explanation, where high ability individuals unable to signal select into self-employment, was described in Section 1.1. However, it is also possible that the returns to education are simply inversely

Table 1.5: *Heckman First Stage: Marginal Effects*

	Self-employed	Employee
Age	0.0121	0.0283
	(0.0002)	(0.0003)
Age^2	-0.0001	-0.0004
	(0.0000)	(0.0000)
Male	0.0803	0.0589
	(0.0005)	(0.0010)
Black	-0.0250	-0.0078
	(0.0007)	(0.0019)
Hispanic	-0.0146	-0.0080
	(0.0009)	(0.0021)
Other Race	0.0103	0.0118
	(0.0013)	(0.0027)
High School	0.0235	-0.0025
	(0.0011)	(0.0019)
Some College	0.0282	-0.0126
	(0.0012)	(0.0019)
College Degree	0.0411	-0.0455
	(0.0015)	(0.0019)
Graduate Degree	0.0720	-0.1022
	(0.0020)	(0.0019)
Married	-0.0063	-0.0153
	(0.0007)	(0.0012)
Disabled	-0.0308	-0.1437
	(0.0006)	(0.0015)
Metropolitan Resident	-0.0084	0.0242
	(0.0006)	(0.0012)
Northeast	-0.0063	0.0062
	(0.0006)	(0.0014)
North Central	-0.0049	0.0135
	(0.0005)	(0.0012)
West	-0.0052	0.0058
	(0.0006)	(0.0014)
Spousal Insurance	0.0664	-0.1371
	(0.0007)	(0.0010)
Log Housing Equity	0.0098	-0.0106
	(0.0002)	(0.0004)
N. of Observations	897,677	897,677

Table 1.6: *Heckman Model: Second Stage*

Log hourly wage	Self-employed	Employee	Difference
Age	0.083***	0.016***	0.067
	(0.005)	(0.001)	
Age^2	-0.001***	-0.000***	-0.001
	0.000	0.000	
Male	0.523***	0.166***	0.357
	(0.015)	(0.004)	
Black	-0.313***	-0.115***	-0.198
	(0.027)	(0.006)	
Hispanic	-0.162***	-0.142***	-0.02
	(0.026)	(0.007)	
Other Race	-0.064**	-0.128***	0.064
	(0.023)	(0.009)	
High School	0.192***	0.110***	0.082***
	(0.023)	(0.006)	(0.022)
Some College	0.355***	0.259***	0.096***
	(0.023)	(0.006)	(0.022)
College Degree	0.556***	0.576***	-0.02
	(0.024)	(0.007)	(0.024)
Graduate	0.992***	0.848***	0.144***
	(0.026)	(0.009)	(0.026)
Married	0.007	0.094***	-0.087
	(0.014)	(0.004)	
Incorporated	0.380***		0.38
	(0.010)		
Business Tenure	0.020***		0.02
	(0.001)		
Job Tenure		0.022***	-0.004
		(0.000)	
Disabled	-0.191***	-0.017	-0.022
	(0.026)	(0.010)	
Metropolitan Area	0.124***	0.161***	-0.174
	(0.011)	(0.004)	
Northeast	-0.037**	0.078***	-0.037
	(0.013)	(0.004)	
North Central	-0.044***	-0.004	-0.115
	(0.012)	(0.004)	
West	0.145***	0.149***	-0.04
	(0.014)	(0.005)	
Inverse Mills	0.277***	-0.351***	0.628
	(0.022)	(0.011)	

Standard errors on difference calculated through bootstrapping.

correlated to traits common among entrepreneurs. Interestingly, self-employed men earn 52.3 percent more than self-employed women. The gender pay differential in the paid sector is only 16.6 percent. This may be capturing educated woman who are staying home with their children and working at home due to situation rather than choice. These women would likely earn lower returns on their education and may be driving the entrepreneurial education penalty that appeared in the OLS results.

1.6.3 Fixed Effects Model

The results from the fixed effects model can be seen in Table 1.7. High school graduates who are self-employed receive a small premium on their degree, but it is statistically insignificant. However, self-employed individuals with some college, a bachelors degree, or a graduate degree see statistically significant higher returns to their education. The premium for a 4-year college degree is 12.1 percent. This is inconsistent with the signaling explanation of education. To test the robustness of these findings, I run a specification of the model that controls for changes in an entrepreneur's business equity. Combining changes in business equity with income to calculate the hourly wage only strengthens the findings of the baseline model. The returns to education become higher for the self-employed across all degree categories. A college degree is worth 20.3 percent more to entrepreneurs than to paid employees. Also, the wage penalty for self-employment diminishes but does not disappear. It falls to only 11.8 percent after controlling for business equity. This indicates that the self-employment wage penalty prevalent in the literature cannot be explained by inaccurate measures of entrepreneurial income.

In the final specification, medical and legal professionals are dropped from the sample. This adjustment reverses the sign on the returns to education for the self-employed with the highest level of education. Those with graduate degrees go from receiving a 10.5 percent premium return on education when they are self-employed to receiving a 16.7 percent penalty on education. It is possible that graduate degrees convey a valuable signal to employees, but not entrepreneurs. But, it is imprudent to make such inferences with any certainty since the occupation restrictions are rather arbitrary and the number of individuals with advanced degrees is relatively small.[22] Other studies that treat education as a continuous variable may be biased by the inclusion of workers with graduate degrees.

Time Variant Selection

The fixed effects model controls for unobserved individual characteristics that are time invariant. However, it cannot control for selection into or out of self-employment that depends on time variant variables. And, the factors that motivate people to enter or leave self-employment are not random. For example, individuals who are laid off may resort to self-employment as an alternative to unemployment. Similarly, leaving a good job voluntarily, is suggestive of an attractive and well-thought-out self-employment opportunity. Movements out of self-employment are also subject to selection problems. Entrepreneurs whose businesses are unsuccessful will exit self-employment and their movement is an indication of failure. However, an exit from self-employment may also indicate the sale of a successful business or the retirement of a business owner. These selection issues can be classified by the direction of the move (into or out of self-employment) and the nature of the move (voluntary or involuntary).

Unfortunately, as these variables tend to vary across individuals but not within individuals, they cannot be estimated within a fixed effects model. Therefore, to test for the presence of time invariant selection issues I run four separate fixed effects regression by type of movement. Individuals who move from self-employment to employment are in one group and are contrasted with those who move the opposite direction. Similarly, individuals who leave their job or business voluntarily are contrasted with those who leave involuntarily. The results can be seen in Table A.1.

First, I examine the differences between voluntary and involuntary movements. The

[22] In an attempt to be less arbitrary, I ran an alternative model excluding all individuals in occupations that are subject to occupational licensing restrictions in all 50 states. The results were very similar to when only doctors and lawyers were excluded.

Table 1.7: *Fixed Effects Models*

Log hourly wage	(a) Base	(b) Equity	(c) No Prof.	(d) No Prof. + Equity
Age	0.105***	0.073***	0.106***	0.077***
	(0.004)	(0.004)	(0.004)	(0.004)
Age^2	−0.001***	0.000	−0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High School x Self	0.058	0.114**	0.058	0.116***
	(0.031)	(0.035)	(0.030)	(0.034)
Some College x Self	0.078*	0.136***	0.080**	0.137***
	(0.031)	(0.035)	(0.031)	(0.035)
College x Self	0.121***	0.203***	0.111***	0.209***
	(0.034)	(0.038)	(0.034)	(0.038)
Graduate x Self	0.105**	0.200***	−0.167***	0.025
	(0.038)	(0.043)	(0.042)	(0.048)
Self-Employed	−0.303***	−0.118***	−0.308***	−0.118***
	(0.025)	(0.029)	(0.025)	(0.029)
Married	−0.003	0.007	−0.006	0.006
	(0.006)	(0.007)	(0.007)	(0.007)
Incorporated	0.017*	0.015	0.005	0.004
	(0.007)	(0.009)	(0.008)	(0.010)
Job Tenure	0.005***	0.003***	0.005***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Business Tenure	0.003***	0.029***	0.005***	0.029***
	(0.001)	(0.001)	(0.001)	(0.001)
Disabled	−0.042***	−0.062***	−0.044***	−0.069***
	(0.006)	(0.007)	(0.006)	(0.007)
Metropolitan Area	0.025**	0.009	0.024**	0.012
	(0.008)	(0.010)	(0.008)	(0.010)
Northeast	0.04	0.066*	0.032	0.056
	(0.026)	(0.029)	(0.026)	(0.029)
North Central	0.043	0.075**	0.049*	0.082**
	(0.023)	(0.026)	(0.023)	(0.026)
West	0.057*	0.072**	0.069**	0.084**
	(0.023)	(0.026)	(0.024)	(0.026)
(%) Unemployed	−0.044***	−0.104***	−0.046***	−0.104***
	(0.007)	(0.008)	(0.008)	(0.008)
Constant	−0.240**	−0.565***	−0.308***	−0.678***
	(0.077)	(0.087)	(0.078)	(0.087)
N. of Observations	414,114	402,759	400,318	389,749

self-employment wage penalty is 16.4 percent larger for individuals who leave their job or business involuntarily compared to those who leave voluntarily.[23] This may be capturing individuals who are laid off from their job or those who reenter the work force after failing at running a small business. However, there is no clear pattern in the returns to education for the self-employed. Many of the coefficients are insignificant. There simply may not be enough underlying variation to identify differences at this level of inquiry.

Next, I examine the impact of the direction of the move. The self-employment wage penalty is 14.2 percent larger for people who move from a job to self-employment relative to those who move in the opposite direction. The interaction of education and self-employment is positive for those who move into self-employment and negative for those who move out of self-employment. This may be capturing that education is less important after a period of job market experience. Again, these signs are largely insignificant, especially for movements out of self-employment.

1.7 Conclusion

Measuring the value of education to entrepreneurs is very valuable. It can help policymakers promote entrepreneurship and it promises to shed light on whether education is valuable because it increases human capital or acts as a signal. However, it is quite difficult to accurately compare the returns to education between entrepreneurs and employees. People not only self-select into self-employment, but education is correlated with innate ability. Most of the literature making inferences about signaling from the returns to education has not addressed these issues sufficiently. I use a stylized model to show how inferences about signaling will only be accurate after controlling for the endogeneity of education and selection into self-employment. To identify the impact of selection on the returns to education I compare a Heckman model to an ordinary least squares regression. This indicates that selection into entrepreneurship biases the returns to education for entrepreneurs downward.

[23] Comparing the coefficient on $Self-employed$ in column (a) to column (b).

To control for all time invariant individual heterogeneity (which deals with the endogenous nature of education) I utilize a fixed effects model. The fixed effects model indicates that a college degree is worth about 12 percent more to the self-employed. Adjusting the earnings of entrepreneurs to compensate for the increasing value of their businesses only increases the education wage premium enjoyed by entrepreneurs. This contradicts the signaling model and goes beyond the human capital model's prediction of equal returns. It appears that education has a greater impact on the productivity of entrepreneurs than employees. This findings should be taken with caution given the methodological difficulties of comparing returns across the two groups. While the fixed effects model solves for time invariant individual heterogeneity, it is possible that time variant selection is influencing the findings. Nevertheless, given that this paper's results are consistent with the other panel study of this issue (Van der Sluis, Van Praag and Vijverberg, 2009) it is worth considering what might be able to explain a premium return on education for entrepreneurs. It seems unlikely that the effect of education on entrepreneurial networks could generate the premium seen in the data. Van der Sluis, Van Praag and Vijverberg (2009) suggest that the self-employed are able to choose work that complements their human capital. But, I find this theory problematic. Employees are also able to pick careers where their human capital is most highly valued. Moreover, high school and college degrees rarely convey specific job skills. Rather, they provide strong foundations in math, English, science and exercise students' broad analytical and communication skills. It may be that being well rounded in these areas is more beneficial to an entrepreneur who must do many different tasks than to an employee who is more likely to specialize on a single task. This would be consistent with the jack-of-all-trades model of entrepreneurship proposed by Lazear (2004). There is ample need for research to confirm and explain the entrepreneurship education premium.

Chapter 2

Do the Unemployed Make Good Entrepreneurs? A Model and Evidence

2.1 Introduction

Policymakers have a longstanding interest in self-employment as an antidote to unemployment. Entrepreneurs[1] create their own jobs and, with any success, create jobs for their employees. For these reasons, many developed countries encourage the unemployed to start their own businesses. For example, Spain, the United Kingdom, France, Sweden and Germany all provide financial incentives for the unemployed to start small businesses. Some of these programs are quite large. For example, Germany spent the equivalent of 680 million USD on its program in 1998 (Reize, 2000). The United States' relatively small program is known as the Self-Employment Assistance Program. Changes to the law in 2012 expanded the scope of this program and relaxed the requirement that participants continue to search for full-time work to remain eligible.[2] Given the policy implications, the relationship between unemployment and self-employment has received considerable attention by economists.

Much of the literature focuses on the macro-level relationship between rates of unemployment and entry into self-employment. In theory, recessions affect participation in en-

[1] Consistent with other work, I will use the terms self-employment and entrepreneurship interchangeably even though many individuals who are self-employed are not particularly innovative.

[2] See the Middle Class Tax Relief and Job Creation Act.

trepreneurship via two channels. According to the "recession-push" hypothesis pioneered by Oxenfeldt (1943), individuals facing low job market prospects have a lower opportunity cost of becoming self-employed and are more likely to enter self-employment. This indicates a positive relationship between unemployment and self-employment rates. In contrast, the "prosperity-pull" hypothesis argues that high rates of unemployment are indicative of recessions. During economic downturns the cost of used capital is higher and the demand for entrepreneurs' products is lower. This makes self-employment less desirable and implies a negative relationship between self-employment and unemployment. Empirical studies conflate these two effects. Parker (2004) provides a comprehensive review of the literature. He finds that most cross-sectional studies report a negative relationship between the local unemployment rate and the probability that an individual is self-employed. In contrast, most time-series studies show a positive relationship between local unemployment rates and entrepreneurship entry. Efforts to reconcile these findings have been of limited success.

While the effect of local unemployment rates on local self-employment rates is unclear, there is strong consensus that unemployed individuals themselves are more likely to enter self-employment.[3] Evans and Jovanovic (1989) find that between 1968 and 1986 unemployed individuals were almost twice as likely to enter self-employment as employees were. Von Greiff (2009) finds that displaced workers are almost twice as likely to enter self-employment in the year after displacement. Similarly, unemployment has been linked to entry into self-employment in other developed countries such as Spain, Canada and the United Kingdom (Carrasco, 1999, Kuhn and Schuetze, 2001, Meager, 1992).

This evidence has raised the concern that individuals who are "pushed" into self-employment may be of lower quality than typical entrepreneurs. It is likely that the unemployed as a group have lower cognitive ability. Heckman, Stixrud and Urzua (2006) find that cognitive skills affect employment outcomes and Heineck (2011), using German data, finds the unem-

[3]For an unemployed individual the "recession-push" effect dominates the "prosperity-pull" effect. But, the unemployed are only a subset of the labor force. Since the "prosperity-pull" affects all the labor force, it is quite possible for the self-employment rates in the general population to fall as the unemployment rate rises.

ployed to have somewhat lower cognitive skills. At the very least, the unemployed tend to be less valuable as workers than typical employees. Therefore, Evans and Jovanovic (1989) label these workers as "misfits." They find that, in addition to the unemployed, individuals receiving relatively low wages and who have changed jobs frequently are most likely to enter entrepreneurship. Von Greiff (2009) argues that during times of high unemployment "misfits" characterized by low-income potential and low levels of wealth dominate entry into self-employment. In contrast, cross-sectional studies indicate that the self-employed typically have higher levels of education and assets (Hamilton, 2000). Moreover, evidence suggests that successful entrepreneurs are well-rounded "jack-of-all trades" (Lazear, 2004). There is concern that public policies that encourage the unemployed to transition into self-employment may be setting up such workers for failure. But, the relevant question is not whether formerly unemployed entrepreneurs are as successful as typical entrepreneurs. Rather, the question of interest is whether displaced workers are better off seeking paid employment or becoming self-employed. It is certainly plausible that "misfits" who may be eccentric, risk-loving, or difficult to work with might have a comparative advantage working for themselves. It would be ideal to compare the earnings of a displaced worker when he becomes self-employed to the earnings he would have received if he had taken paid employment. Since this is impossible, I employ an alternative approach.

I utilize a simple model to characterize the duration of self-employment spells. I show that in equilibrium, survival rates should be unrelated to an entrant's innate ability. While low ability workers will earn less as entrepreneurs, they will also earn less in paid employment. Therefore, the probability an individual will survive in self-employment depends on how his earnings compare to his expectations. However, a growing number of studies using international data find that unemployed individuals who transition into self-employment do not survive as long as former employees. Taylor (1999), using British data, finds that each previous month of unemployment reduces the probability of survival by 1.2 percent. Andersson and Wadensj (2007), employing a probit model, find that formerly unemployed Swedes are 11-18 percent more likely to exit self-employment. Using Spanish data, Carrasco (1999) finds that unemployment increases the hazard rate in the first few quarters by between two and three fold. For simplicity, I will refer to this negative relationship between former unemployment and self-employment survival as the duration puzzle. To my knowledge, there have been no U.S. studies on the duration puzzle.

To explain the duration puzzle, I suggest that unemployed individuals overestimate their entrepreneurial potential at entry. I put forth this hypothesis on the basis of evidence from behavioral economics and the psychology literature.

A consistent finding in behavioral economics is that individuals frequently overestimate their abilities. One popular example is that 93 percent of American drivers consider themselves to be more skillful than a median driver (Svenson, 1981). Another study finds that MBA students systematically overestimate the number of job offers they will receive (Hoch, 1985). Similarly, economists find that the levels of trading volume in financial markets far exceed what a model based on rational investment would predict. Barber and Odean (2001) provide evidence that investor overconfidence is responsible. The pattern that emerges over a number of studies is that overconfidence is greatest for tasks that are difficult, unpredictable, and lacking in meaningful feedback (Barber and Odean, 2001). These traits characterize the decision an individual faces when evaluating whether to enter self-employment. The market is often unfamiliar and the variation in possible outcomes is high. Therefore, it is not surprising that empirical work finds entrepreneurs overconfident in their abilities. In a study of 2,994 recent entrants to self-employment, 81 percent perceived their odds of success to be better than 70 percent (Cooper, Woo and Dunkelberg, 1988). An outstanding 33 percent felt that their success was guaranteed. Of particular interest, individuals with poor backgrounds (low amounts of education and experience etc.) were just as optimistic about their chances of success.

To my knowledge, no study has tested whether overconfidence is related to an individual's work status prior to entry. However, there are compelling reasons to suspect that unemployed workers transitioning into self-employment might be especially prone to overconfidence. As noted earlier, unemployed workers as a group have below average ability. Recent developments in the psychology literature find that low ability individuals are particularly prone to overconfidence. As explained in the first of many articles on the subject by Kruger and Dunning (1989) (pg. 1121):

> "In essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain–one's own or anyone else's. Because of this, incompetent individuals lack what cognitive psychologists...term metacognition...the ability to know how well one is performing, when one is likely to be accurate in judgment, and when one is likely to be in error."

Experimental tests by the authors in the fields of humor, grammar, and logic show that participants in the bottom quartile grossly overestimate their performance while participants in higher quartiles have much more accurate assessments of their performance. An informal way of expressing this finding is that because low ability individuals "don't know what they don't know," they are overconfident.

Additionally, job displacement itself may contribute to overconfidence. Workers have imperfect information about ability and therefore use wages to make inferences about their true ability. However, wages are a function not only of ability but of firm rents and firm-specific human capital. Bheim, Horvath and Winter-Ebmer (2010) find evidence that workers overattribute high wages in the past to innate ability. Workers do not realistically account for their firm-specific value when making inferences about their market value. As a result, they find that workers displaced from companies paying wages with large firm-specific components are unemployed longer than usual. The implication is that displaced workers imperfectly interpret their past wages and therefore have an inflated view of their innate ability. This may make available job offers seem unattractive relative to self-employment. The problem is exacerbated by the tendency of layoffs to target workers with unjustifiably high wages.

Also, the effects of job loss and/or job searching have unique psychological effects. The unemployed face social stigma and financial strain. Winkelmann and Winkelmann (1998) find that the non-pecuniary costs of unemployment far exceed the monetary costs of forgone wages. A meta-analysis of over 300 studies on the mental health effects of unemployment finds that unemployment is linked to lower self-esteem and higher rates of depression (Paul and Moser, 2009). These symptoms were particularly pronounced for extended periods of unemployment. It is therefore likely that as the duration of unemployment prior to entry into self-employment increases, an individual's confidence at entry will decline. These individuals, therefore, should be less likely to exit self-employment. This feature acts as another testable implication of overconfidence.

It should be noted that the effect of unemployment on overconfidence will depend on how a person became unemployed. Individuals who are laid-off will have a much different psychological profile from those who take advantage of early retirement. Some unemployed individuals may be taking a leisurely break from work while others will be struggling to make ends meet. Unfortunately, data limitations make it difficult to know why an individual is unemployed. It is possible to make generalization about the unemployed as a group. However, any public policy efforts should acknowledge that there are substantial differences between different types of unemployed people.

2.2 Theoretical Model

2.2.1 Background

An influential paper by Jovanovic (1982) models entrepreneurs as heterogeneous individuals with imperfect information about their abilities. His model implies that transitions into and out of self-employment are driven by changing beliefs about individuals' innate entrepreneurial ability. This was a major deviation from previous models such as Lucas (1978) who assumes that individuals know their own ability and Kihlstrom and Laffont (1979) who assume that entrepreneurship is driven by risk preferences. Jovanovic (1982) is now the standard model in the literature as its predictions are consistent with a large body of empirical evidence (Parker, 2004). My model is based on Frank (1988), an extension of Jovanovic (1982).[4] While Frank (1988) allows workers to learn about their ability more quickly by working more, I require hours to be constant across all individuals. Unlike Frank (1988), I solve the model numerically to generate a number of predictions about self-employment duration. Later, I tweak the model to allow workers to overestimate their true ability.

2.2.2 Setup

A risk-neutral individual is considering self-employment. Upon entry, the firm will be a price taker. Each period the entrepreneur will decide to keep operating ($b_t = 0$) or exit ($b_t = 1$). If entrepreneurship is unsuccessful, the individual can return to the traditional labor market and earn wage U_t (in terms of utility). The outside wage, U_t, does not change as long as the worker is self-employed.[5] In order to enter self-employment, the individual must incur a sunk cost C. There are no costs to shut down. There are a finite number of periods T. Individuals do not know their own entrepreneurial ability. However, the distribution of entrepreneurial ability, α, is known to be normally distributed such that $\alpha \sim \mathcal{N}(m_1, \sigma_\alpha^2)$. Earnings, R_t, are noisy signals of ability. It is known to individuals that the process which generates R_t is:

$$R_t = \alpha + \epsilon_t \tag{2.1}$$

where ϵ_t is known to be normally distributed such that $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$. At entry, an entrepreneur's best guess of his ability is the population average m_1.[6] As the entrepreneur remains in the market and observes R_t, he uses Bayes' rule to update the estimated value of α. The updated guess is denoted by m_{t+1}. This estimate will also be normally distributed

[4] I base my model on Frank (1988) because it can be solved via backwards induction and is more concise than Jovanovic (1982).

[5] In other words, job search and full-time entrepreneurship are mutually exclusive. Also, while U_t is constant through time, the time subscript is included to make later comparisons with time variant variables more clear.

[6] It is not necessary for all individuals in the economy to be identical. Therefore, m_1 can be thought of the population average conditional on observed characteristics such as age, education etc.

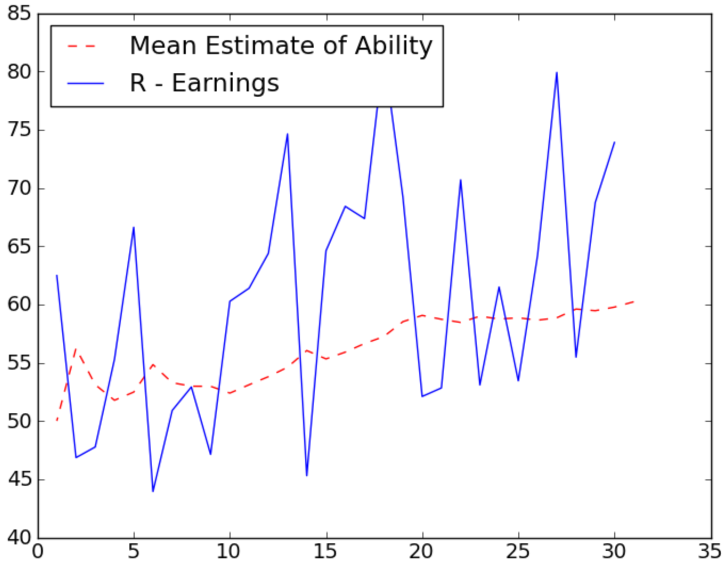
with mean m_{t+1} and precision (inverse of variance) h_{t+1} as follows[7]:

$$m_{t+1} = \frac{h_t m_t + h_\epsilon R_t}{h_t + h_\epsilon} \tag{2.2}$$

$$h_{t+1} = h_t + h_\epsilon \tag{2.3}$$

where $h_\epsilon = \frac{1}{\sigma_\epsilon^2}$. As an entrepreneur remains in the market, precision increases and the estimate of ability becomes more accurate. To illustrate this process, consider a new self-employed person with innate entrepreneurial ability of 60. In period 1, the individual's best guess of his ability is 50, the population average. Through time the individual's earnings will allow him to make a more informed estimate of ability using Equation 2.2. This process is illustrated in Figure 2.1.[8]

Figure 2.1: *Updating Estimates of Entrepreneurial Ability*



Let Π_t be the present discounted value of an individual's expected utility at time t if he acts optimally henceforth. E_t is the expectation operator based on the information at time

[7]This is proven in DeGroot (2005). I am also able to confirm the predictions of the formula using Bayesian inference.

[8]The individual knows the distribution of α and ϵ. In this example, $\sigma_\epsilon = 10$ and $\sigma_\alpha = 10$.

t. Each period, the entrepreneur observes earnings R_t which are discounted by β where $(0 < \beta \leq 1)$. The individual faces the following maximization problem:

$$\Pi_1 = \max_{b_1} \{b_1 U_1 + (1 - b_1) [E_1 R_1 - C] + \beta E_1 \Pi_2\} \\ t = 1 \tag{2.4}$$

$$\Pi_t = \max_{b_t} \{b_t U_t + (1 - b_t) E_t R_t + \beta E_t \Pi_{t+1}\} \\ t = 2, 3 ... T \tag{2.5}$$

Each period, the entrepreneur chooses to exit $(b_t = 1)$ if the outside wage is greater than the expected earnings today plus the net present value of having the option of entrepreneurship tomorrow. Part of the incentive to remain self-employed is that entrepreneurial ability may turn out to be higher than currently thought (the individual having been just unlucky so far). It is worthwhile to learn more about one's entrepreneurial ability because if one is indeed talented at entrepreneurship the payoff can be quite high. Once an individual exits self-employment, re-entry will not occur, as m_t will remain constant and U_t is time invariant.

2.2.3 Solution

The model captured by Equations 2.1-2.5 can be solved via backwards induction. In the final period, there is no future so there is no value in remaining self-employed to learn about unknown entrepreneurial ability. Therefore, the maximization problem in the final period reduces to:

$$\Pi_T = \max_{b_T} \{b_T U_T + (1 - b_T) m_T\} \tag{2.6}$$

The individual will remain self-employed $(b_T = 0)$ if his estimated entrepreneurial ability exceeds the outside wage $(m_T > U_T)$. Otherwise, the entrepreneur will exit $(b_T = 1)$. In the second to last period, the entrepreneur faces the following maximization problem:

$$\Pi_{T-1} = \max_{b_{T-1}} \{b_{T-1} U_{T-1} + (1 - b_{T-1}) m_{T-1} + \beta E_{T-1} \Pi_T\} \tag{2.7}$$

where:

$$E_{T-1}\Pi_T = b_{T-1}U_T + E_{T-1}\left[b_T|b_{T-1}=0\right]U_T + \left(1 - E_{T-1}\left[b_T|b_{T-1}=0\right]\right)E_{T-1}\left[m_T|b_T=0\right] \tag{2.8}$$

If the individual exits in $T-1$ ($b_{T-1}=1$), then by definition, he will not be an entrepreneur in period T. Therefore, the individual will receive U_T in the final period. If the individual does not exit in $T-1$, he will exit in period T if the reservation wage in period T exceeds his best guess of entrepreneurial ability such that $U_T > m_T$. However, m_T is unknown at time $T-1$. The best guess of ability next period is simply the current estimate of ability, therefore $E_{T-1}\left[m_T|(b_{T-1}=0)\right] = m_{T-1}$. The probability of exit in the terminal period is calculated as shown:

$$\mathrm{E}_{T-1}\left[b_T|b_{T-1}=0\right] = \Phi\left(\frac{U_T - m_{T-1}}{1/\sqrt{h_T}}\right) \tag{2.9}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function. Since precision increases by a constant amount each period, h_T is known at $T-1$.

Moreover, if the entrepreneur does not exit in period T, it is likely that the entrepreneur received favorable revenue in $T-1$. This indicates that entrepreneurial ability is high and the expectation of revenue in the final period should be high. The updated expectation can be calculated using Bayesian inference. Let $\pi(\theta)$ represent the prior distribution such that $\pi(\theta) \sim \mathcal{N}(m_{T-1}, \frac{1}{h_{T-1}})$. The new information, $b_T = 0$, implies that $m_T \geq U_T$.[9] This new information is denoted by E. The probability of observing E for a given value of θ is represented by $l(E|\theta)$. The posterior distribution is formed as follows:

$$f(\theta|E) = \frac{\pi(\theta)l(E|\theta)}{\int \pi(\theta)l(E|\theta)d\theta} \tag{2.10}$$

I approximate this discretely with the help of a Python package from Downey (2011). The mean of the posterior distribution is $\mathrm{E}_{T-1}\left[m_T|b_T=0\right]$.

Given a set of base parameters, it is possible to numerically solve for the optimal choice of b_{T-1}. By simulating the model using many different values of of m_{T-1}, I am able to

[9]This is only true in the final period. Since, there is no value in acquire more information about unobserved ability the maximization problem is simplified.

calculate the minimum value of m_{T-1} for which the entrepreneur will stay self-employed denoted by m_{T-1}^{min}. For convenience, I will refer to this as the ability threshold. Using backwards induction, it is possible to solve for m_t^{min} all the way back to the first period. I assume that entrants into self-employment do so at the margin. In other words, entrants have an m_1 that is just slightly larger than m_1^{min}. This assumption is later relaxed and explored further.

After working backwards to solve for all m_t^{min}, I work forward to calculate the survival function of marginal entrants. By design, for all entrants, $m_1 > m_1^{min}$. Therefore, the probability of exit for all individuals is 0 in the first period. In the first period, some entrants receive low earnings and revise their new estimate of ability, m_2, downward. Probability of exit in period two is as shown:

$$p(b_2 = 1) = \Phi\left(\frac{m_2^{min} - m_2}{1/\sqrt{h_2}}\right) \tag{2.11}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function. Finally, entrepreneurs who survive period two are of higher ability than average. This means that the average ability of entrepreneurs, m_t, is increasing through time. Conversely, the probability of exit, conditional on surviving thus far, is falling through time. The probability of exiting in period three is given by:

$$p(b_3 = 1) = \Phi\left(\frac{m_3^{min} - E(m_3|b_2 = 0)}{1/\sqrt{h_3}}\right) \tag{2.12}$$

Bayesian inference is used to calculate, $E(m_3|b_2 = 1)$, which accounts for the impact of survival on mean ability. This equation can be generalized to solve for the probability of exit in $t = 4 \ldots$T. Given the exit probabilities, it is straightforward to generate the population cumulative survival function. The survival functions generated by the model will later be compared visually to survival functions generated by the data.

To quantitatively make comparisons between the theoretical and empirical models it is useful to generate hazard rates. The hazard rate, h_t, is the probability an individual will exit in period t, conditional on surviving through period $t-1$. It is the complement of

Table 2.1: *Baseline Parameters*

Parameter	Description	Value
U	Wage possible in paid employment	50.0
β	Discount factor	.95
T	Number of periods	20
σ_α	Standard deviation of ability	69
σ_ϵ	Standard deviation of earnings noise	51
C	First period sunk cost	5

the slope of the survival function. Let p_t^{cs} denote the cumulative probability of survival in period t. The hazard rate in period t is defined as:

$$h(t) = \frac{p_{t-1}^{cs} - p_t^{cs}}{p_{t-1}^{cs}} \tag{2.13}$$

By comparing the hazard rates of two models that differ by only one variable, it is possible to quantify the marginal impact of a variable on exit. These comparisons are known as hazard ratios.

2.2.4 Baseline Model Predictions

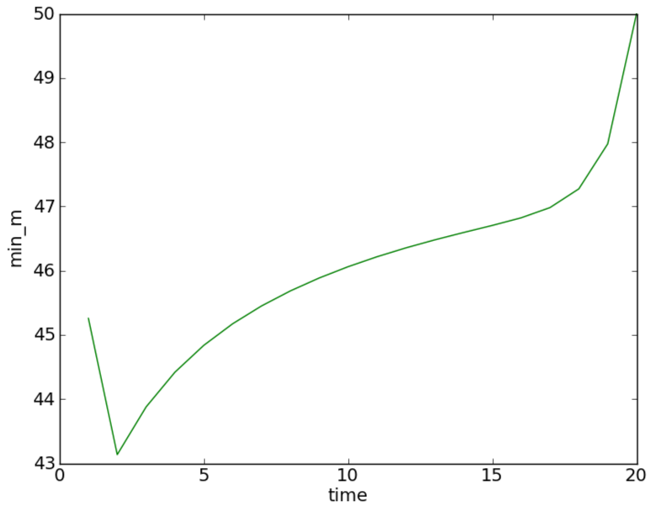
The model is not calibrated to match key statistics. Rather, it is used to gain intuition about the behavior of entrepreneurs and to generate testable predictions of this behavior. To make these predictions, it is necessary to start with some baseline parameters. The parameters in the baseline model and their chosen values are shown in Table 2.1. The discount factor (β), the number of periods (T) and the first period sunk costs (C) do not affect the dynamics of the model and are chosen arbitrarily.[10] Entry and exit decisions depend primarily on the standard deviation of entrepreneurial ability (σ_α) and the standard deviation of earnings noise (σ_ϵ). These parameters are based on empirical summary statistics from the 1996 Survey of Income and Program Participation (SIPP). The outside wage is arbitrarily chosen to be 50 and the standard deviations are normalized relative to it.

The minimum estimate of ability (m_t^{min}) that an individual must have to rationally

[10] I considered using data on business capital to inform the measure of sunk costs, but it is impossible to know what portion of capital investments are sunk.

continue in self-employment is graphed through time in Figure 2.2. The m_1^{min} to enter self-employment is 45.26. Since the outside wage, U, equals 50, this means that individuals will knowingly enter self-employment even though they expect their self-employment earnings to be less, on average, than what they would have earned in paid-employment.[11] This is rational because the information individuals obtain about their entrepreneurial ability is very valuable. If entrepreneurial ability is revealed to be low, an individual can cut his losses incurring only the sunk cost of entry and the opportunity cost of having forgone more lucrative paid work. However, if entrepreneurial ability is revealed to be high, the individual will enjoy earnings higher than U for the rest of his career.

Figure 2.2: *Baseline Model: Minimum Required Estimate of Ability*



After entry, m_t^{min} drops precipitously. Because the sunk costs of entry have been incurred, the marginal cost of seeing another period of earnings information is now lower. In the subsequent periods, the m^{min} function is initially concave capturing the diminishing marginal informational value of another period of earnings data. However, a countervailing force eventually causes the m_t^{min} function to become convex. As time passes and the

[11]This could help explain the self-employment wage penalty that is consistently found in the literature. A cross-sectional snapshot will make it appear as though the self-employed are earning less. However, this would not necessarily be the case if the researcher was able to view lifetime earnings.

individual nears the terminal period, the value of information about entrepreneurial ability declines. In the final period, T, there is no value to learning about entrepreneurial ability because there is no future. Therefore, $m_T^{min} = U$. This force makes the m_t^{min} function convex because the percentage decline in the future is increasing in magnitude over time.[12]

Figure 2.3: *Baseline Model: Probability of Surviving Through Period t*

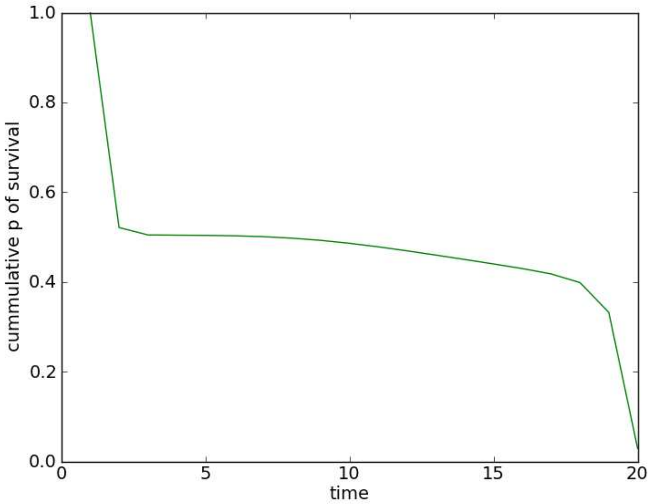


Figure 2.3 shows the probability an entrant will survive through period t. As before, entrants are marginal such that $m_1 \approx m_1^{min} \approx 45.26$. By definition, all entrants survive the first period and the probability of survival strictly falls through time. In the baseline model, almost half of the entrants exit after seeing one period of earnings data. This steep decline is driven by the parameters σ_ϵ and σ_α. The larger the variation within an individual's earnings, the less informative earnings become, decreasing the likelihood of exit. The higher the variation in earnings across entrepreneurs the lower the average ability level of entrants. Since entrants are marginal, their average entrepreneurial ability is below the population average. It is therefore not surprising that most entrants lose the "entrepreneurial lottery." Informative earnings allow entrants to recognize this quickly. The surviving entrepreneurs have higher ability and the marginal informational value of new earnings is decreasing with time.

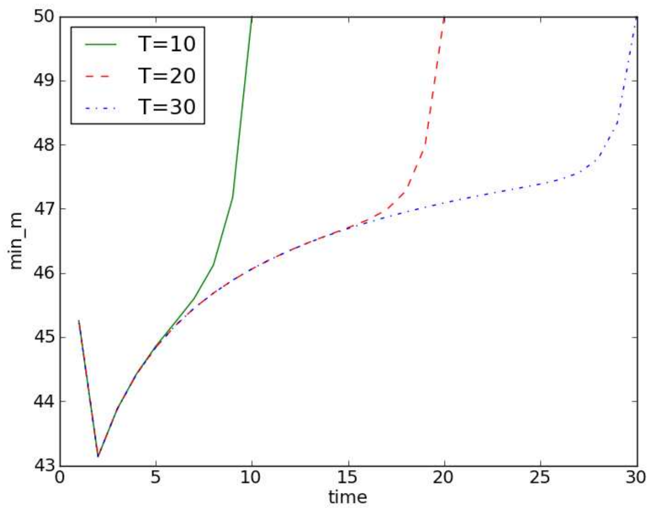
[12]From t=16 to t=17 the number of periods in the future goes from 4 to 3, a 25% reduction. From t=17 to t=18 the number of periods in the future goes from 3 to 2, a 33.3% reduction.

This causes the survival function to flatten rather rapidly. As the entrepreneur approaches the final period, the ability threshold rises rapidly and the survival rate plummets. Empirically, this effect will be less pronounced as entrepreneurs have a career that is indefinite rather than finite in length.

2.2.5 Variations to the Baseline Model

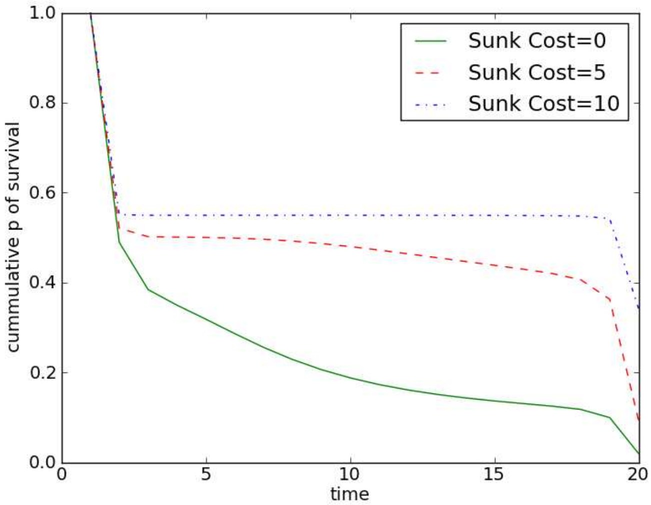
While the baseline model is quite informative, further insights are possible by varying the parameters. In particular, I explore the consequences of varying the time frame (T), the sunk costs (C), the variation of entrepreneurial ability in the population (σ_α) and the variation in the error component of earnings (σ_ϵ). The implications are as follows:

Figure 2.4: *Baseline Model: m_t^{min} as T varies*



1. As the time frame increases, the ability threshold, m_t^{min}, falls. This is because a longer future increases the lifetime value of being self-employed if entrepreneurial ability is revealed to be high. This is illustrated in Figure 2.4. It should be noted that in the first few periods the minimum ability threshold seems to be nearly identical in all specifications. The curves only diverge as t increases. Intuitively, when the future consists of just a few periods, the potential entrant must consider both whether he has

Figure 2.5: *Relationship Between Sunk Costs and Survival*

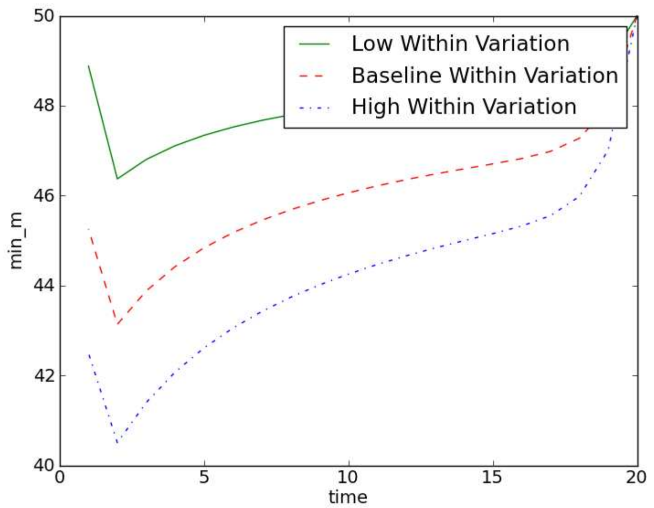


a comparative advantage in entrepreneurship and whether the future is long enough to recoup the cost of entry. As T increases, the problem quickly simplifies to whether the individual expects entrepreneurial ability to exceed the outside wage, expressed as $E_t(\alpha > U)$.

2. As first-period sunk costs increase, m_1^{min} must rise. This is because it becomes more costly for low-ability workers to enter self-employment as a long-shot gamble. However, because the costs are sunk, they have no effect on m_t^{min} after entry. Sunk costs do, however, increase the average ability of entrants. Therefore sunk costs are positively correlated with survival rates. At extremely high levels of sunk costs, only very high ability workers will enter the market and subsequent exits will be virtually non-existent. This is illustrated in Figure 2.5.

3. As the standard deviation of the error component of earnings, σ_ϵ, increases, m_t^{min} falls. This is illustrated in Figure 2.6. This happens because as σ_ϵ increases, earnings become less informative indicators of ability. Therefore, an individual's estimate of ability has less precision. With perfect precision, an individual will only enter self-employment if

Figure 2.6: *Relationship of Within Variation and Ability Threshold*



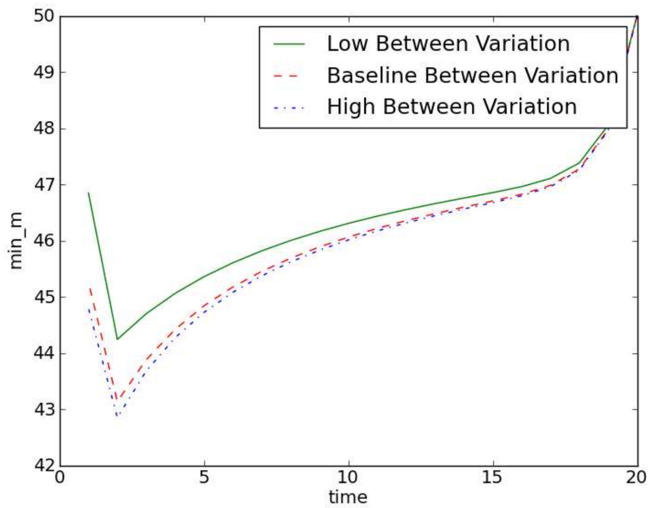
$m_1 > 50$, that is, if self-employment pays more than the outside wage. However, as guesses of ability become less precise, low estimates of ability are increasingly likely to be due to chance.

4. As the standard deviation of entrepreneurial ability in the population, σ_α, increases, m_t^{min} must fall. This is true because as σ_α rises, the probability that a given individual has exceptional entrepreneur ability increases. It is also true that the probability a given individual will fail miserably increases. However, these entrepreneurs can cut their losses quickly and go back to paid employment. Successful entrepreneurs will enjoy a premium income until period T. A high σ_α, increases the potential gains of entrepreneurship without increasing the potential losses. Therefore, σ_α is inversely correlated with m_t^{min}. This is shown in Figure 2.7 where the baseline level of σ_α is varied by 50 percent in either direction.

2.2.6 Duration Puzzle

As noted before, there is empirical evidence that the unemployed have shorter self-employment durations than average. The goal of the model is to provide a theoretical explanation for

Figure 2.7: *Relationship of Between Variation and Ability Threshold*

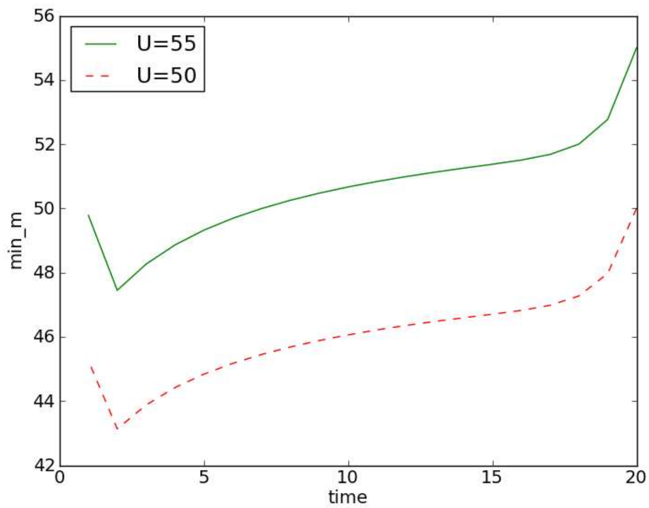


this puzzle. I begin by ruling out two possible explanations.

Differences in innate ability cannot explain the duration puzzle. At first glance, it might appear that differences in ability are responsible for the duration puzzle. It is quite plausible that the unemployed are, on average, of lower ability than employed workers. In this case, it seems natural that lower ability individuals would fail more quickly at self-employment. However, low ability individuals also would earn a lower outside wage in the traditional labor market. In fact, the entrepreneurial ability of marginal entrants is derived in the model from the outside wage. The impact of these two opposing forces should cancel each other out.

Consider a deviation from the baseline model where two types of workers are considering self-employment. Low ability workers have an outside wage $U = 50$. Given the parameters of the baseline model, marginal low ability entrants will have a mean estimated innate ability of $m_1 = 45.26$. High ability workers have outside wage that is 10 percent higher, so $U = 55$. Given the parameters of the baseline model, marginal high ability entrants will have a mean estimated innate ability of $m_1 = 50.26$. The impact on m_t^{min} is shown in Figure

Figure 2.8: *Relationship Between Outside Wage and Ability Threshold*



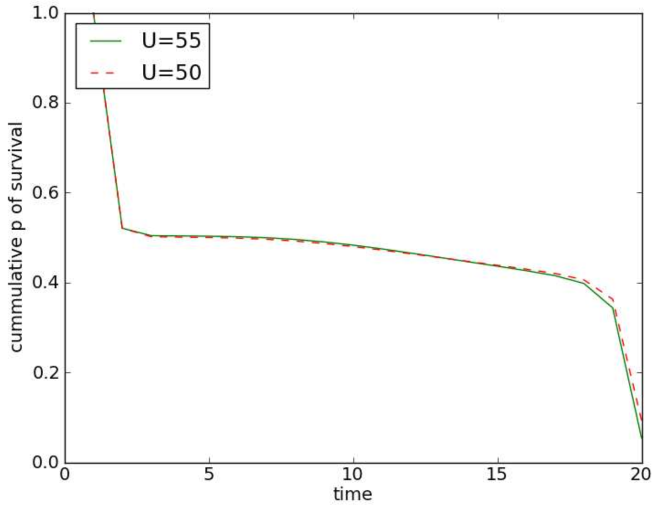
2.8. High ability individuals face a higher m_t^{min} than low ability individuals. However, high ability entrants have higher entrepreneurial ability and earn higher entrepreneurial earnings. Therefore, survival rates are identical between the two groups as shown in Figure 2.9.[13]

Relaxing the assumption of marginal entry cannot explain the duration puzzle. One of the assumptions of the model is that entrants to self-employment do so at the margin. This means that entrants are indifferent between entering self-employment and receiving the outside wage. The justification for this assumption is that any entrants for whom entry was particularly attractive would already have entered in prior periods. However, in the case of the unemployed, this assumption is less tenable. Some subset of the unemployed who enter self-employment do so after receiving a shock to their outside wage, U. A negative shock to U in the last period could induce an individual to enter self-employment. For this individual, it is likely that $m_1 > m_1^{min}$.

To explore the consequences of this possibility, it is useful to compare the survival rates of workers who do and do not enter at the margin. Consider the baseline model where

[13] This is true after scaling up the variation and sunk cost parameters by 10 percent. Remaining differences in the survival functions are merely the result of rounding errors in the discrete approximation.

Figure 2.9: *Relationship Between Outside Wage and Survival*

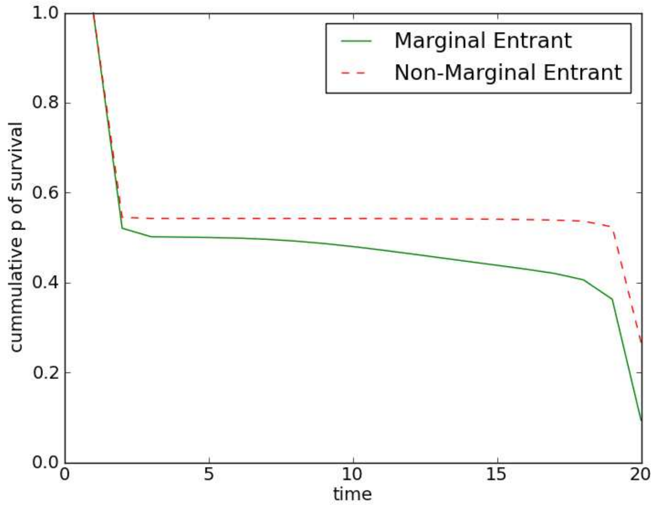


normal workers have an outside wage of 50 and enter at the margin. This implies that $m_1 = m_1^{min} = 45.26$. In addition, suppose that a group of unemployed workers received a wage of 55 last period. Now that they have lost their jobs, the outside wage they could receive if they take a new job this period is only 50. For these workers, $m_1 > [m_1^{min}(U = 50) = 45.26]$. However, since these workers did not enter self-employment last period, it is known that $m_1 < [m_0^{min}(U = 55) = 50.26]$. So, for the marginal entrant $m_1 = 45.26$ but for the non-marginal entrant $45.26 < m_1 < 50.26$. Relaxing the assumption of marginal entry increases the average ability of entrants while leaving the ability threshold unchanged. Therefore, non-marginal entrants should have even higher survival rates, contrary to the duration puzzle. This is illustrated in Figure 2.10 when the mean estimate of ability for non-marginal entrants is chosen to be $m_1^{min} = 47.76$ (the midpoint between $m_1^{min}(U = 50) = 45.26$ and $m_1^{min}(U = 55) = 50.26$).

2.2.7 Overconfidence

The model shows that movements into and out of self-employment are driven by imperfect information. In a world with perfect information no one would exit self-employment.

Figure 2.10: *Relationship Between Marginal Entry and Survival*



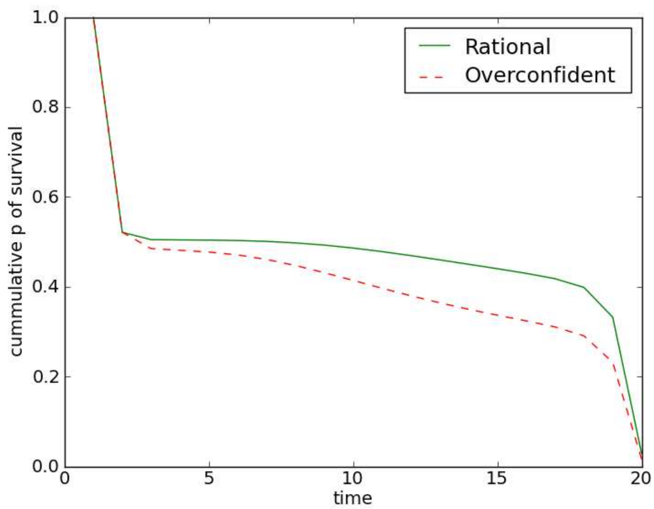
Individuals with low entrepreneurial ability would never enter and those with high entrepreneurial ability would never leave. Survival rates, therefore, depend primarily on the evolution of information over time. Up until this point, I have assumed that individuals have rational expectations about their ability. In other words, individuals have accurate information about the distribution of entrepreneurial ability and earnings. However, if individuals are systematically biased in their expectations, it will substantially affect survival rates. In particular, individuals who overestimate their entrepreneurial ability will be more likely to exit self-employment.

To illustrate this, consider the baseline model where the outside wage is 50. Suppose there are two groups: rational individuals and overconfident individuals. Rational individuals have an accurate assessment of the distribution of their entrepreneurial ability. Overconfident individuals believe that their mean entrepreneurial ability is 1 percent higher than what a rational person would believe. Therefore, an overconfident individual with the background to justify an expected entrepreneurial ability of 50 would actually believe his expected entrepreneurial ability to be 50.5. In the model, this overconfidence is captured by $\gamma = 1.01$ where $m_1 = \gamma E_1[\alpha]$. Overconfidence only affects initial estimates of ability.

Overconfident individuals still rationally update their estimate of ability after seeing their entrepreneurial earnings.

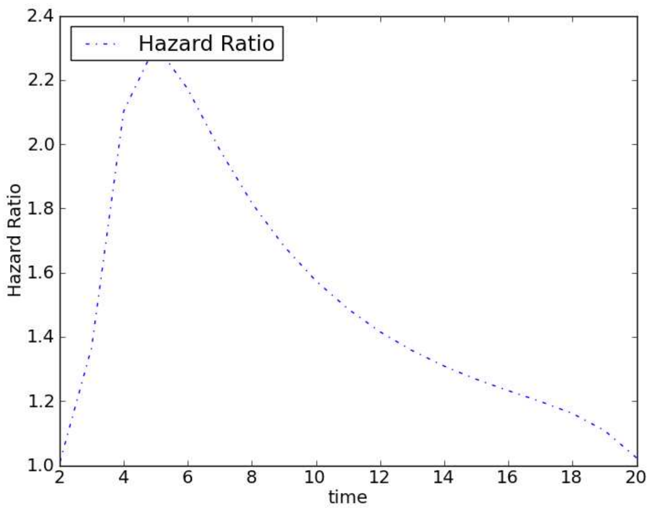
Since the two groups have the same U, σ_α, and σ_ϵ both groups will have the same ability threshold through time. In the first period, $m_1^{min} = 45.26$. For rational individuals, $E_1[\alpha]$ must exceed 45.26 to make self-employment attractive. For an overconfident individual, $\gamma E_1[\alpha]$ must exceed 45.26. In the case of $\gamma = 1.01$ an $E_1[\alpha]$ of 44.82 would be sufficient for entry. The effect of overconfidence is to decrease the true ability of marginal entrants while leaving the perceived ability constant across both groups. Following entry, individuals in the overconfident group will be more likely to receive lower than expected earnings. This will cause them to revise their estimates of ability downward and a disproportional number of them will exit self-employment. This is illustrated in Figure 2.11.

Figure 2.11: *Relationship Between Overconfidence and Survival*



It is quite clear from Figure 2.11 that overconfident individuals are more likely to exit entrepreneurship. What is less apparent visually is that the effect of overconfidence on survival is changing over time. To see this, it is useful to compare the hazard rate of overconfident individuals to the hazard rate of rational individuals. The hazard rates are calculated as described in Equation 2.13. Since overconfident individuals are of lower ability, they tend to see lower earnings than rational individuals. The more periods of low earnings they see, the more likely they will be to revise their estimates of ability downward enough to exit the market. This causes the hazard rate to increase with time. Eventually, however, all the overconfident individuals who should have exited will have already done so. In other words, overconfidence makes marginal workers more likely to exit, but it does not have much of an effect on high ability workers. Together, these two forces result in a concave hazard ratio as shown in Figure 2.12. Amazingly after 5 months, individuals in the baseline model that overestimate their ability by just 1 percent are twice as likely to exit self-employment.

Figure 2.12: *Hazard Rate of Overconfident/Hazard Rate of Rational*



2.3 Empirical Work

In this section, I replicate the duration puzzle finding that the formerly unemployed are more likely to exit self-employment. Then, I provide some preliminary evidence to support the hypothesis that overconfidence is responsible for the duration puzzle.

2.3.1 Data

To achieve these aims, I use pooled data from the 1996 and 2001 panels of the Survey of Income and Program Participation (SIPP). The SIPP is well-suited to study transitions of the unemployed into self-employment. Most importantly, the SIPP surveys respondents quarterly and asks them to provide information about each of the last 4 months. This results in monthly observations, a critical property since the average duration of self-employment is quite short. Another advantage of the SIPP is its large size. In the 1996 SIPP, over 40,000 households are interviewed quarterly from 1996-2000. Individuals in the 2001 SIPP are interviewed from 2000-2003. Together, this generates over 5 million monthly observations of working-age individuals. Because unemployment and self-employment entry are relatively rare, a large data set is necessary for statistically significant estimates. The main disadvantage of the SIPP is that it does not contain much background information on respondents. It would be ideal to have measures of innate ability such as IQ in order to compare the ability of those individuals who enter self-employment from unemployment to those who do not. In the absence of this information, I use education as an imperfect proxy of ability.

The SIPP specifically asks respondents about their employment status. This makes it possible to accurately determine spells of joblessness without having to make inferences from less reliable variables such as income. Official statistics require individuals to be actively seeking work to be considered unemployed. I do not make this distinction for two reasons. While the SIPP collects some data on job searching, relative to macroeconomic statistics, it appears to be missing a large number of individuals looking for work. Moreover, I'm only studying individuals who move from unemployment to self-employment. These individuals are revealing themselves to be members of the labor force. Whether they were actively searching for jobs prior to entering self-employment is of little consequence. To focus on individuals in the labor market, I exclude minors and people attending school.

I generate a dummy to capture whether an individual was unemployed in the month prior to self-employment. In addition, I create the variable $TimeUnemployed$ to measure

Table 2.2: *Differences Between Entering and Incumbent Entrepreneurs*

	Entrant	Incumbent	Difference	T-Value
Education (years)	13.60	13.90	-0.30	-24.66
High School (%)	0.29	0.29	0.00	-0.19
Some College (%)	0.29	0.27	0.02	9.08
College (%)	0.19	0.20	-0.01	-8.47
Graduate (%)	0.12	0.15	-0.03	-17.43
Married (%)	0.73	0.76	-0.03	-14.95
Age (years)	46.59	46.51	0.08	1.62
Male (%)	0.66	0.74	-0.08	-41.29
Black (%)	0.05	0.04	0.01	14.11
Hispanic (%)	0.08	0.06	0.02	20.54
Incorporated (%)	0.26	0.39	-0.12	-63.05
Monthly Income ($)	$3,276.97	$4,192.18	-$915.21	-$38.78
Hourly Wage ($)	$23.37	$25.86	-$2.50	-$7.59
Weekly Hours	43.05	47.54	-4.49	-60.24

the cumulative time an individual has been unemployed. After, a spell of employment or self-employment, cumulative unemployment is reset to zero.

2.3.2 Summary Statistics

Since I am interested in the job market status of individuals prior to their self-employment spell, I restrict the duration analysis to individuals who are observed to enter self-employment during the panel. This means an individual who enters the panel self-employed and leaves the panel self-employed would not be observed in my analysis. Due to selection effects, entrants to self-employment are of lower quality then incumbents. Entrants tend to have less education and earn lower incomes. The differences between the two groups of entrepreneurs are shown in Table 2.2. Because of these differences, summary statistics about self-employed entrants should not be considered to be representative of the self-employed as a whole. Rather, the following results are suggestive for an important slice of entrepreneurs: those who are just starting out. In this panel, entrants compose 60.1 percent of all self-employment observations.

The theoretical model shows that self-employment duration is a function of variation in

Table 2.3: *Variation in Income During Self-employment*

	Unemployed at Entry	Typical Entrant	All Entrants
Mean Income	$2,333	$3,548	$3,106
Overall Standard Deviation	3,840	5,531	5,016
Within Standard Deviation	2,047	3,097	2,761
Between Standard Deviation	3,225	4,284	3,897
# of Individuals	2,092	2,824	4,910

ability across individuals and within an individual's earnings. These variables are approximated by the within and between standard deviations of income as shown in Table 2.3.[14] Not surprisingly, individuals who enter self-employment from unemployment have lower average earnings. However, relative to income, the within and between variation are of the same magnitude for both groups. This is encouraging as it indicates that neither differences in the distribution of ability nor the level of earnings noise between groups are responsible for the duration puzzle.

When individuals exit self-employment, they are asked to report the reason for their departure. Of the 5,788 entrants to self-employment observed during the panel, 4,584 exited and reported the reason for leaving self-employment. Their answers are shown in Table 2.4. About 12 percent of exits were the result of business failure. That is not much higher than the 9.1 percent that retired. In contrast, 25.4 percent of respondents quit to take another job.[15] These results are consistent with the theoretical model where self-employment exit is driven by the opportunity cost of being self-employed and is not modeling business failure per se.

Of the 5,788 entrants to self-employment, 2,531 are unemployed prior to entry. Consistent with the literature, the unemployed are more likely to enter self-employment. On average, about 1.4 percent of individuals who are not currently self-employed will enter

[14] Formerly unemployed and typical individuals do not sum to the total for "All" because a few individuals have multiple spells of unemployment and fit in both categories. Also, income totals do not match summary statistics totals because some individuals have missing values for income.

[15] The actual question in the SIPP is quit to "To start other business/take job," but I restrict respondents to those who are no longer self employed.

Table 2.4: *Reasons for Leaving Self-employment*

Reason	# Responses	% of Responses
Retirement or old age	419	9.1%
Childcare problems	42	0.9%
Other family/personal problems	359	7.8%
Own illness	192	4.2%
Own injury	55	1.2%
School/training	29	0.6%
Went bankrupt/business failed	550	12.0%
Sold business or transferred	271	5.9%
To take a job	1,164	25.4%
Season ended for a seasonal	204	4.5%
Quit for some other reason	1,299	28.3%
Total	4,584	100.0%

self-employment in the next 12 months. Using a probit model I find that unemployed workers are 26.2 percent more likely to enter self-employment. The full results are available in Appendix B.1. This finding suggests that unemployed workers may not be entering at the margin. Which, as noted in the theory section, only makes their shorter survival times in self-employment more puzzling.

Finally, it is informative to contrast the differences between the unemployed and employed as they transition into self-employment. Theory predicts that entrants with lower outside wages will have lower innate ability. While the SIPP does not contain a measure of innate ability, education at entry and income during duration are informative signals of innate ability. The differences in ability and other background variables at entry can be seen in Table 2.5. Entrants transitioning from unemployment have less education, are more likely to be female, are less likely to be married, and are more likely to be a minority. The formerly unemployed are 2.3 years younger, on average, at entry.

Following entry, the formerly unemployed underperform relative to former employees. These differences are shown in Table 2.6. The formerly unemployed bring home 962 fewer dollars each month relative to their employee counterparts. Only 19 percent of their businesses are incorporated compared to 28 percent for former employees. They also work fewer

Table 2.5: *Differences at Entry by Prior Job Status*

	Unemployed	Employed	Difference	T-Value
Education (years)	13.20	13.80	-0.59	-32.56
High School (%)	0.32	0.29	0.03	11.84
Some College (%)	0.28	0.28	0.00	-0.37
College (%)	0.16	0.20	-0.04	-15.59
Graduate School (%)	0.10	0.14	-0.04	-21.84
Married (%)	0.67	0.75	-0.08	-28.74
Age (years)	44.51	46.85	-2.34	-28.15
Male (%)	0.60	0.71	-0.11	-37.61
Black (%)	0.08	0.04	0.03	20.65
Hispanic (%)	0.13	0.07	0.06	31.58
Individuals Entering	2,531	3,257		

Table 2.6: *Differences in Self-employment Outcomes by Prior Job Status*

	Unemployed	Employed	Difference	T-Value
Monthly Income ($)	$2,509.06	$3,470.83	-$961.78	-31.72
Hourly Wage ($)	$22.57	$23.57	-$1.00	-1.85
Weekly Hours	38.48	44.20	-5.72	-46.17
Incorporated (%)	0.19	0.28	-0.10	-35.57
Average Duration (months)	8.0	9.4	-1.47	
Individuals Entering	2,531	3,257		

hours each week. These results are not particularly surprising. Since the unemployed are of lower ability they should earn lower wages. However, according to the predictions of the model, they should not exit self-employment prematurely. Yet, they do. The average duration of self-employment is 8.0 months for individuals who were previously self-employed and 9.4 months for regular workers.[16] Consistent with the literature, it appears that the formerly unemployed do not survive as long in self-employment. To show that this difference is statistically significant and explore whether it can be explained by overconfidence will require a more formal approach.

2.3.3 Methodology

While the average duration of self-employment is informative, it is useful to see how survival rates vary through time. To accomplish this, I employ the Kaplan-Meier Curve. Let $S(t)$ be the probability that a self-employed entrant will survive to period t. In each time period t_i the number of "at risk" individuals is denoted by n_i and the number of exits (or deaths) is denoted by d_i. The Kaplan-Meier estimator is the nonparametric maximum likelihood estimate of $S(t)$ as follows:

$$\hat{S}(t) = \prod_{t_i<t} \frac{n_i - d_i}{n_i} \tag{2.14}$$

One of the benefits of the Kaplan-Meier Curve is that it is not biased by censoring as n_i simply diminishes if individuals drop out of the panel. To illustrate the duration puzzle, I create separate Kaplan-Meier curves for the formerly unemployed and employed and compare them using the log-rank test.

It is likely that variables correlated to unemployment, rather than unemployment itself, are influencing survival. These variables are denoted by vector X. To isolate the effect of unemployment on survival, I use the Cox proportional hazards model (Cox, 1972). As with all hazard models, the hazard rate, $h(t, X)$, gives the probability of exit in period t

[16]These durations may seem short. That is because they are downward biased due to censoring when the panel ends (after 4 or 3 years depending on the panel).

conditional on surviving so far. For a discrete duration random variable T:

$$h(t) = \frac{Pr(T = t)}{Pr(T \geq t)} \tag{2.15}$$

In a proportional hazards model, it is assumed that part of the hazard rate is constant across all individuals and is solely a function of time. This is referred to as the underlying hazard function, $\lambda(t)$. It captures how the hazard changes over time as other variables are held constant. This causes the effect of variables to be multiplicatively related to the underlying hazard function. For example, if being male increases the rate of hazard by 10 percent in the first period, it should also increase the rate of hazard by 10 percent in the second period and all subsequent periods. The distinctive feature of the Cox model is that the effect of variables on the hazard rate can be estimated without having to specify the underlying hazard function. For this reason, the Cox model is considered to be a semiparametric model. For a vector of variables X, the hazard function for the Cox model is given by:

$$h(t, X) = \lambda(t)e^{X'\beta} \tag{2.16}$$

The parameters of the model can be estimated using partial maximum likelihood estimation. Let t_i denote the time period entrepreneur i is observed to exit given that t_i is one of r observed exit times $\{t_{(1)}, t_{(2)}, \cdots, t_{(r)}\}$. Self-employment duration is right censored for individuals who are still self-employed when the survey ends. The partial likelihood function is:

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{e^{X_i'\beta}}{\sum_{t_j \geq t_i} e^{X_j'\beta}} \right]^{\delta_i} \tag{2.17}$$

where $\delta_i = 0$ if i is right censored and $\delta_i = 1$ otherwise.

The log of the above function is maximized using standard maximum likelihood techniques to estimate the parameters. The model is solved using the Stata package stcox and ties are handled using Breslow's method (Breslow, 1975).

I run two specifications of the model. First, I seek to confirm the duration puzzle by showing that unemployment is responsible for shorter self-employment durations after controlling for all possible variables. To achieve this aim, the chosen variables in the first specification of the model are as follows:

$$X'\beta = Demographics'\beta_1 + Education'\beta_2 + \beta_3 Unemployed \tag{2.18}$$

Demographics includes age, gender, race, marital status, and region. *Education* includes dummies for high school through graduate school. *Unemployed* is a dummy equal to 1 if an individual was unemployed prior to entering self-employment. If the unemployed exit self-employment early then $\beta_3 < 0$.

The second objective is to test the hypothesis that the self-employed are overconfident. It would be ideal to ask the unemployed and employed at entry their expected chances of success and test the accuracy of their predictions. Lacking this data, I look for evidence consistent with overconfidence by the unemployed. The model simulations predict that the hazard ratio of unemployment should be time dependent. In the simulated model, the hazard ratio initially increases with time before hitting a maximum and decreasing. I test for this phenomena by interacting self-employment duration with a dummy for formerly unemployed. I also include the square of this variable.

Additionally, I test whether the length of unemployment duration affects the hazard rate in self-employment. In other words, are people who are unemployed for a long period of time less likely to exit self-employment? If overconfidence is responsible for the duration puzzle then there are a few reasons that unemployment duration might reduce confidence. As unemployment duration increases, the literature indicates that people become less optimistic and have reduced self-esteem. The psychological costs of unemployment might counteract initial overconfidence. Moreover, as time goes on the unemployed will search for jobs, receive job offers, and receive better feedback about their value in the marketplace. This may give them a clearer view of their innate ability. As a result, it should reduce initial biases formed if the unemployed do not accurately recognize the firm-specific component of past wages when making inferences about their current market value. I test both of these predictions using the following specification:

$$
\begin{aligned}
X'\beta = Demographics'\beta_1 + Education'\beta_2 + \beta_3 Unemployed \\
+\beta_4 TimeUnemployed + \beta_5 TimeUnemployed^2 \\
+\beta_6 Unemployed * SelfDuration + \beta_7 (Unemployed * SelfDuration)^2
\end{aligned} \tag{2.19}
$$

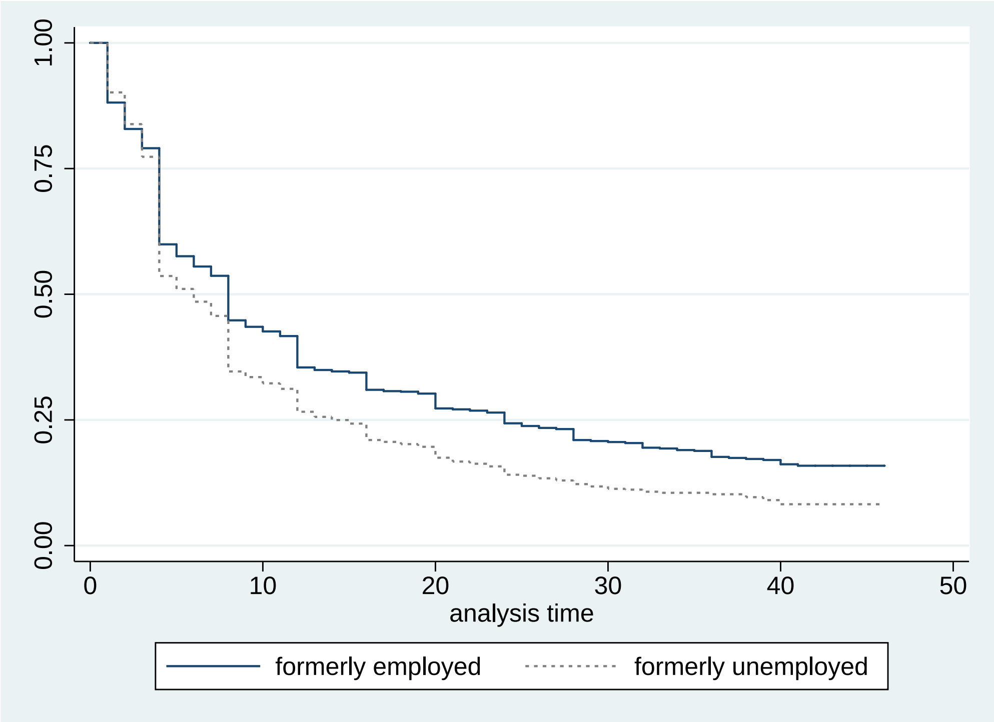
$SelfDuration$ is the amount of time the entrant has survived in self-employment. If individuals eventually become less overconfident the longer they are unemployed then $\beta_4 < 0$. If self-employment duration affects overconfidence in a concave fashion as described earlier then $\beta_6 > 0$ and $\beta_7 < 0$.

2.3.4 Results

The Kaplan-Meir survival curves for both groups are shown in Figure 2.13. Visually, the shape of both survival curves is consistent with the theoretical model. Survival probability initially drops sharply before leveling off. In terms of magnitude, slightly fewer entrepreneurs survive empirically than theory predicts. But, that is not surprising given the admittedly rough estimates of parameters. Most importantly, the survival rate of the formerly unemployed is much lower than that of the formerly employed. Over time, the survival rates appear to diverge at a decreasing rate. Using the log-rank test, the difference between the two curves is statistically significant at a 99 percent confidence level.

The results from both specifications of the Cox model are shown in Table 2.7. To aid in interpretation, the estimates are shown both as coefficients and hazard ratios. The hazard ratio is the exponent of the estimated coefficient. It is interpreted as the hazard rate if the variable increases by one unit, divided by the baseline hazard rate. For example, the hazard ratio on $Male$ is .80 in the second specification. This means that the hazard rate of someone who is male is .80 percent the size of the hazard rate of a female. So, men are less likely to exit self-employment than women. The effects of the variables are largely consistent across the two specifications. Older people, men, and those living in metropolitan areas are less likely to exit self-employment. College completion does not have a statistically significant

Figure 2.13: *Kaplan-Meier Self-employment Survival Estimates*



impact on the hazard rate. People in Western states are more likely to exit entrepreneurship. This may be because there is less stigma associated with business failure in the West.

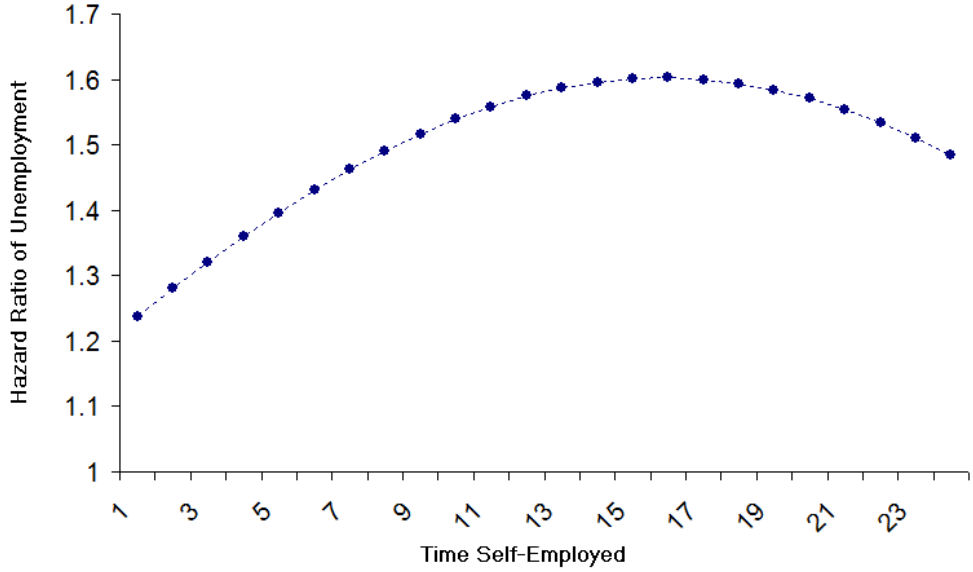
The first specification confirms the duration paradox. After controlling for other variables, the unemployed are 27.5 percent more likely to exit from self-employment. This effect is statistically significant at the 99 percent level. This confirms that, as in many other European countries, unemployed workers in the United States are less likely to survive in self-employment.

The second specification shows that the effect of unemployment on the hazard ratio initially increases, peaks and then decreases. Self-employment duration amplifies the hazard rate of unemployment while self-employment duration squared decreases the hazard rate. To interpret the net effect of these forces, I calculate the combined impact on the hazard ratio and graph the results in Figure 2.14. The hazard ratio peaks at 1.6 after 18 months of self-employment. This means that after 18 months of being self-employed, formerly unemployed workers are 60 percent more likely to exit self-employment than an otherwise

Table 2.7: *Results from the Cox Model*

	Specification 1		Specification 2	
	Hazard Ratio (p-value)	Coefficient (p-value)	Hazard Ratio (p-value)	Coefficient (p-value)
Age	0.955	-0.046	0.954	-0.048
	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.807	-0.215	0.799	-0.224
	(0.000)	(0.000)	(0.000)	(0.000)
Black	0.948	-0.053	0.954	-0.047
	(0.386)	(0.386)	(0.445)	(0.445)
Hispanic	0.904	-0.101	0.908	-0.096
	(0.043)	(0.043)	(0.056)	(0.056)
High School	0.877	-0.132	0.878	-0.130
	(0.008)	(0.008)	(0.009)	(0.009)
Some College	0.948	-0.054	0.946	-0.055
	(0.285)	(0.285)	(0.274)	(0.274)
College	1.045	0.044	1.045	0.044
	(0.431)	(0.431)	(0.428)	(0.428)
Graduate	0.965	-0.036	0.963	-0.038
	(0.585)	(0.585)	(0.563)	(0.563)
Married	1.007	0.007	1.005	0.005
	(0.826)	(0.826)	(0.874)	(0.874)
Metro	0.896	-0.110	0.898	-0.108
	(0.004)	(0.004)	(0.005)	(0.005)
North East	0.954	-0.047	0.956	-0.045
	(0.330)	(0.330)	(0.353)	(0.353)
North Central	1.026	0.026	1.027	0.027
	(0.538)	(0.538)	(0.522)	(0.522)
West	1.129	0.121	1.129	0.122
	(0.003)	(0.003)	(0.003)	(0.003)
Unemployed Dummy	1.275	0.243	1.194	0.178
	(0.000)	(0.000)	(0.018)	(0.018)
Time Unemployed			0.982	-0.018
			(0.043)	(0.043)
$(\text{Time Unemployed})^2$			1.000	0.000
			(0.123)	(0.123)
Unemployed X Self Duration			1.038	0.037
			(0.007)	(0.007)
$(\text{Unemployed X Self Duration})^2$			0.999	-0.001
			(0.024)	(0.024)
N. of Observations	44,447	44,447	44,447	44,447
N. of Subjects	5,387	5,387	5,387	5,387
N. of Failures	3,589	3,589	3,589	3,589
Time at Risk	46,644	46,644	46,644	46,644

Figure 2.14: *Effect of Self-employment Duration on the Unemployment Hazard Ratio*



identical individual who was employed prior to entry. The evolution of the hazard ratio over time is consistent with the behavior of the overconfident individuals in the theoretical model. While not conclusive, this provides some evidence that overconfidence is responsible for the diminished survival of the unemployed.

The specification also shows that individuals unemployed for longer periods of time are more likely to survive in unemployment. Each period of unemployment duration reduces the hazard rate by 1.2 percent. However, for most unemployed individuals, this negative effect is dominated by the positive effect of the unemployment dummy variable for all but the longest unemployment durations. Someone would have to be unemployed for over 2 years before the net effect of unemployment on hazard was negative. This result is consistent with long-term unemployment reducing irrational exuberance and helping individuals to gain a better understanding of their value in the marketplace.

2.4 Robustness

The Cox model is a semiparametric model because it leaves the baseline hazard function, $h(t)$, unspecified. This is appropriate since there is not an a priori reason to expect the hazard function to be of a certain functional form.[17] However, to check the robustness of my results, I compare my findings against two parametric proportional hazards models that specify the functional form of $h(t)$. In the Weibull proportional hazards model, the hazard function is specified as:

$$h(t) = \alpha\lambda^{\alpha}t^{\alpha-1}$$

If $\alpha > 1$ then $h(t)$ is monotonically increasing through time. If $\alpha < 1$ then $h(t)$ is monotonically decreasing through time. In the special case where $\alpha = 1$, the hazard function reduces to $h(t) = \lambda$. In this case, $h(t)$ is constant through time and this model is known as the exponential proportional hazards model.

I rerun Specification 2 using each of these proportional hazards models. The results are shown in Appendix B.2. The alternative models find even stronger evidence of the duration puzzle. In Specification 2 of the Cox model, the unemployed are 19.4 percent more likely to exit self-employment. Whereas in the Weibull model, the formerly unemployed are 2.1 times more likely to exit self-employment and in the exponential model they are 1.9 times more likely to exit. This indicates that the original analysis, if anything, understates the size of the duration puzzle. In contrast, the demographic variables are quite consistent across the three models. It appears that the effect of prior unemployment on the hazard rate may be related to time in ways that the other variables are not. In addition, the original analysis indicates that individuals who enter self-employment after a long period of unemployment are less likely to exit self-employment. This result is robust to both alternative models, changing only slightly in magnitude.

Finally, the original analysis finds the effect of prior unemployment on hazard to vary

[17] In contrast, in medical applications this may not be true. A medical researcher may expect a drug to impact survival during a certain period. For example, a drug may not have any effect if it is administered too early or after a disease has progressed too far.

with time as shown in Figure 2.14. This result is not fully confirmed by the alternative models. While the coefficient on $UnemployedXSelfDuration$ is positive in the Cox model, it is negative in the Weibull model and insignificant in the exponential model. The effects of these differences are graphed relative the original model in Appendix B.1. In the alternative specifications, the effect of unemployment on duration is strictly decreasing with time. In contrast, in the Cox model, the effect of unemployment on duration initially increases and then decreases. The alternative models are not showing evidence that overconfident individuals are, initially, more likely to exit after receiving additional periods of (likely) disappointing earnings. All three models reflect that eventually the effect of prior unemployment on hazard diminishes. Since the alternative parametric models require the imposition of arbitrary restrictions on the baseline hazard function, I argue that the Cox model is more reflective of the data. Nevertheless, these alternative specifications show stronger support for the duration puzzle but somewhat weaker circumstantial evidence consistent with overconfidence by the formerly unemployed.

2.5 Conclusion

There is a growing body of evidence in European countries that the unemployed who transition into self-employment do not survive as long as typical entrants. I replicate this finding using a large United States dataset. Formerly unemployed Americans are 27.5 percent more likely to exit self-employment in a given period. I explore the implications of this result using a theoretical model. The model shows that differences in ability between the unemployed and employees cannot explain differences in self-employment duration. Instead, I suggest that overconfidence in ability by the unemployed may be responsible for the duration puzzle. This hypothesis is informed by evidence in the literature that entrepreneurs as a group tend to overestimate their abilities. Moreover, work in psychology indicates that low-ability individuals lack the metacognition to accurately predict their own performance. Since the unemployed as a group are of lower ability, they should be more prone to overestimate their entrepreneurial potential.

While it is impossible to prove that the unemployed overestimate their ability, I find circumstantial evidence consistent with this theory. The model predicts that the hazard ratio of overconfident to rational individuals should vary with the time someone is self-employed, resembling a parabola in shape. This reflects two opposing forces: 1) Overconfident individuals will be more likely to exit as they receive additional periods of earnings generated from a distribution of ability that is lower than they realized at entry and 2) As time passes, low ability workers will exit leaving only high ability workers who are not at the margin of exiting and for whom overconfidence is inconsequential. This pattern in the hazard ratio is confirmed empirically using a Cox proportional hazards model. The model shows that less than 1 percent overconfidence on the part of the unemployed could generate the substantial hazard ratio seen in the data. Furthermore, I find that the hazard of prior unemployment is smaller for people who have been unemployed for a long time. I argue that this is consistent with the psychological aspects of unemployment.

Entrepreneurship and self-employment are often lauded for stimulating economic growth and generating jobs. Therefore, it is appealing to look to self-employment as a means to transition displaced workers into working, productive members of society. This would be ideal if the unemployed were "misfits" better suited to working for themselves. However, this paper's findings suggest that governments should be cautious before providing incentives for the unemployed to enter self-employment. In fact, if the unemployed overestimate their entrepreneurial ability, they should be discouraged from entering self-employment. Unnecessary spells of self-employment incur sunk costs, reduce work experience and result in foregone wages. Therefore, subsidizing transitions from unemployment to self-employment could reduce both individual and societal welfare.

At this point, more research is needed to better understand why the unemployed fail at entrepreneurship and whether overconfidence is responsible. The stigma of joblessness may be an alternative explanation for why the unemployed have shorter self-employment spells. Additionally, field work in this area would be highly valuable to better understand the motivations and psychology of the unemployed.

Bibliography

Aldrich, Howard E., and Catherine Zimmer. 1986. "The Art and Science of Entrepreneurship." *In: Sexton, D.L., Smilor, R.W. (Eds.).*

Andersson, Pernilla, and Eskil Wadensj. 2007. "Do the unemployed become successful entrepreneurs?" *International Journal of Manpower*, 28(7): 604–626.

Barber, Brad M., and Terrance Odean. 2001. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment." *The Quarterly Journal of Economics*, 116(1): pp. 261–292.

Becker, Gary S. 1975. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, 2nd ed. NBER Books*, National Bureau of Economic Research, Inc.

Bheim, Ren, Gerard Thomas Horvath, and Rudolf Winter-Ebmer. 2010. "Great Expectations: Past Wages and Unemployment Durations." Department of Economics, Johannes Kepler University Linz, Austria Economics working papers 2010-09.

Blanchflower, David G, and Andrew J Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics*, 16(1): 26–60.

Breslow, N. E. 1975. "Analysis of Survival Data under the Proportional Hazards Model." *International Statistical Review / Revue Internationale de Statistique*, 43(1): pp. 45–57.

Brown, Sarah, and John G Sessions. 1998. "Education, Employment Status and Earnings: A Comparative Test of the Strong Screening Hypothesis." *Scottish Journal of Political Economy*, 45(5): 586–91.

Card, David. 1999. "The causal effect of education on earnings." In *Handbook of Labor Economics*. Vol. 3 of *Handbook of Labor Economics*, , ed. O. Ashenfelter and D. Card, Chapter 30, 1801–1863. Elsevier.

Carrasco, Raquel. 1999. "Transitions to and from Self-employment in Spain: An Empirical Analysis." *Oxford Bulletin of Economics and Statistics*, 61(3): 315–41.

Cooper, A.C., T. Folta, and C. Woo. 1991. "Information acquisition and performance by start-up firms." *N.S. Churchill (Ed.) et al., Frontiers of Entrepreneurship Research, Center for Entrepreneurial Studies*, 276– 290.

Cooper, Arnold C., Carolyn Y. Woo, and William C. Dunkelberg. 1988. "Entrepreneurs' perceived chances for success." *Journal of Business Venturing*, 3(2): 97–108.

Cox, D. R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2): pp. 187–220.

DeGroot, M.H. 2005. *Optimal Statistical Decisions. Wiley Classics Library*, John Wiley & Sons.

Downey, A. 2011. *Think Stats.* O'Reilly Media.

Evans, David S., and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *The Journal of Political Economy*, 97(4): 808–82.

Frank, Murray Z. 1988. "An Intertemporal Model of Industrial Exit." *The Quarterly Journal of Economics*, 103(2): 333–44.

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

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–61.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): pp. 411–482.

Heineck, Guido. 2011. "Do your brains help you out of unemployment?" *Economics Letters*, 111(1): 34 – 36.

Heywood, John, and Xiangdong Wei. 2004. "Education and Signaling: Evidence from a Highly Competitive Labor Market." *Education Economics*, 12(1): 1–16.

Hoang, Ha, and Bostjan Antoncic. 2003. "Network-based research in entrepreneurship: A critical review." *Journal of Business Venturing*, 18(2): 165 – 187.

Hoch, Stephen J. 1985. "Counterfactual reasoning and accuracy in predicting personal events." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4): 719–731.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica*, 50(3): 649–70.

Kihlstrom, Richard E., and Jean-Jacques Laffont. 1979. "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion." *Journal of Political Economy*, 87(4): pp. 719–748.

Kruger, Justin, and David Dunning. 1989. "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments." *Journal of Personality and Social Psychology*, 77(6): 1121–1134.

Kuhn, Peter J., and Herb J. Schuetze. 2001. "Self-employment dynamics and self-employment trends: a study of Canadian men and women, 1982-1998." *Canadian Journal of Economics*, 34(3): 760–784.

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

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

Meager, Nigel. 1992. "Does unemployment lead to self-employment?" *Small Business Economics*, 4: 87–103. 10.1007/BF00389850.

Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings. NBER Books*, National Bureau of Economic Research, Inc.

Oxenfeldt, A.R. 1943. *New firms and free enterprise: pre-war and post-war aspects.* American Council on Public Affairs.

Parker, Simon C. 2004. *The Economics of Entrepreneurship. Cambridge Books*, Cambridge University Press.

Paul, Karsten I., and Klaus Moser. 2009. "Unemployment impairs mental health: Meta-analyses." *Journal of Vocational Behavior*, 74(3): 264 – 282.

Reize, Frank. 2000. "Leaving unemployment for self-employment: a discrete duration analysis of determinants and stability of self-employment among former unemployed."

Riley, John G. 1979. "Noncooperative Equilibrium and Market signaling." *American Economic Review*, 69(2): 303–07.

Robb, Alicia, Reedy E.J., Janice Ballou, David DesRoches, Frank Potter, and Zhanyun Zhao. 2008. "An Overview of the Kauffman Firm Survey 20042008." The Kauffman Foundation Overview.

Spence, A. Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics*, 87(3): 355–74.

Spence, Michael. 2002. "Signaling in Retrospect and the Informational Structure of Markets." *American Economic Review*, 92(3): 434–459.

Stiglitz, Joseph, and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review*, 71(3): 393–410.

Svenson, Ola. 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica*, 47(2): 143 – 148.

Taylor, Mark P. 1999. "Survival of the Fittest? An Analysis of Self-Employment Duration in Britain." *The Economic Journal*, 109(454): 140–155.

Van der Sluis, Justin, Mirjam Van Praag, and Wim Vijverberg. 2008. "Education And Entrepreneurship Selection And Performance: A Review Of The Empirical Literature." *Journal of Economic Surveys*, 22(5): 795–841.

Van der Sluis, Justin, Mirjam Van Praag, and Wim Vijverberg. 2009. "Returns for Entrepreneurs vs. Employees: The Effect of Education and Personal Control on the Relative Performance of Entrepreneurs vs. Wage Employees." Institute for the Study of Labor (IZA) IZA Discussion Papers 4628.

Van Praag, Mirjam, and Peter H. Versloot. 2008. "The Economic Benefits and Costs of Entrepreneurship: A Review of the Research." *Foundations and Trends in Entrepreneurship*, 4(2): 65–154.

Von Greiff, Jenny. 2009. "Displacement and self-employment entry." *Labour Economics*, 16(5): 556–565.

Vos, Ed, Andy Jia-Yuh Yeh, Sara Carter, and Stephen Tagg. 2007. "The happy story of small business financing." *Journal of Banking & Finance*, 31(9): 2648–2672.

Weiss, Andrew. 1995. "Human Capital vs. Signaling Explanations of Wages." *Journal of Economic Perspectives*, 9(4): 133–54.

Winkelmann, Liliana, and Rainer Winkelmann. 1998. “Why Are the Unemployed So Unhappy? Evidence from Panel Data.” *Economica*, 65(257): 1–15.

Wolpin, Kenneth. 1977. “Education and Screening.” *American Economic Review*, 67(5): 949–58.

Appendix A

Chapter 1 Supporting Documents

Table A.1: *Fixed Effects Models by Type of Move*

Log hourly wage	(a) Voluntary	(b) Involuntary	(c) Into Self	(d) Out of Self
	Coef./se	Coef./se	Coef./se	Coef./se
Age	0.125***	0.148***	0.127***	0.042
	(0.009)	(0.009)	(0.032)	(0.030)
Age^2	-0.001***	-0.001***	-0.001**	0
	(0.000)	(0.000)	(0.000)	(0.000)
High School x Self	-0.01	0.147**	-0.116*	0.154**
	(0.046)	(0.053)	(0.058)	(0.054)
Some College x Self	-0.136**	0.259***	0.032	0
	(0.047)	(0.056)	(0.058)	(0.057)
College x Self	0.097	-0.016	-0.039	0.175**
	(0.049)	(0.064)	(0.064)	(0.059)
Graduate x Self	-0.063	0.013	-0.102	0.11
	(0.059)	(0.081)	(0.077)	(0.067)
Self-employed	-0.238***	-0.402***	-0.186***	-0.328***
	(0.039)	(0.044)	(0.050)	(0.046)
Married	-0.043**	-0.011	-0.296***	-0.178***
	(0.014)	(0.015)	(0.049)	(0.046)
Incorporated	0.057*	0.332***	0.111***	0.060*
	(0.023)	(0.030)	(0.031)	(0.030)
Job Tenure	0.003***	0.013***	0.002	0.007**
	(0.001)	(0.001)	(0.003)	(0.002)
Business Tenure	0.005**	-0.008**	0.004	0.009**
	(0.002)	(0.002)	(0.002)	(0.003)
Disabled	-0.046**	-0.082***	-0.208***	-0.084
	(0.015)	(0.012)	(0.047)	(0.045)
Metropolitan Area	0.036*	0.007	0.159*	0.159*
	(0.017)	(0.018)	(0.063)	(0.067)
Northeast	0.105**	-0.039	-0.374	-0.658**
	(0.038)	(0.059)	(0.295)	(0.243)
North Central	0.062	0.062	-0.063	0.242
	(0.034)	(0.050)	(0.162)	(0.136)
West	0.141***	-0.022	0.426*	-0.415*
	(0.036)	(0.046)	(0.184)	(0.205)
Constant	-0.960***	-1.549***	-0.912	1.354*
	(0.183)	(0.192)	(0.701)	(0.624)
N. of Observations	67,506	62,888	10,855	12,126

Into Self are individuals who move from a job to self-employment.

Out of Self are individuals who move in the opposite direction.

Appendix B

Chapter 2 Supporting Documents

Figure B.1: *Alternative Models: The Effect of Self-employment Duration on the Unemployment Hazard Ratio*

Table B.1: *Probit Model: Entry into Self-Employment*

	Marginal Effect	Standard Errors	%Δ in Entry Probability
Age	0.0014	0.0002	0.0976
Age^2	0.0000	0.0000	-0.0010
Black	-0.0057	0.0007	-0.4096
Hispanic	0.0006	0.0010	0.0397
Other	-0.0005	0.0012	-0.0354
High school	0.0012	0.0009	0.0858
Some college	0.0025	0.0010	0.1788
College	0.0040	0.0012	0.2827
Graduate	0.0061	0.0016	0.4346
Male	0.0081	0.0006	0.5799
Married	0.0038	0.0006	0.2707
Disabled	-0.0037	0.0007	-0.2638
Metro	-0.0010	0.0006	-0.0746
North East	-0.0022	0.0007	-0.1565
North Central	-0.0015	0.0007	-0.1100
West	0.0010	0.0007	0.0682
Unemployed	0.0037	0.0011	0.2619
Observations	1,413,182		

Table B.2: *Alternative Models for Specification 2*

	Cox	Weibull	Exponential
	Hazard Ratio (p-value)	Hazard Ratio (p-value)	Hazard Ratio (p-value)
Age	0.954	0.939	0.941
	(0.000)	(0.000)	(0.000)
Male	0.799	0.763	0.766
	(0.000)	(0.000)	(0.000)
Black	0.954	0.972	0.972
	(0.445)	(0.695)	(0.684)
Hispanic	0.908	0.930	0.929
	(0.056)	(0.209)	(0.197)
Other	1.176	1.206	1.206
	(0.018)	(0.025)	(0.022)
High School	0.878	0.865	0.866
	(0.009)	(0.013)	(0.012)
Some College	0.946	0.928	0.930
	(0.274)	(0.212)	(0.214)
College	1.045	1.031	1.031
	(0.428)	(0.640)	(0.635)
Graduate	0.963	0.950	0.951
	(0.563)	(0.505)	(0.510)
Married	1.005	1.004	1.006
	(0.874)	(0.914)	(0.887)
Metro	0.898	0.873	0.875
	(0.005)	(0.003)	(0.003)
North East	0.956	0.930	0.934
	(0.353)	(0.209)	(0.224)
North Central	1.027	1.027	1.027
	(0.522)	(0.594)	(0.592)
West	1.129	1.143	1.140
	(0.003)	(0.006)	(0.006)
Unemployed Dummy	1.194	2.100	1.886
	(0.018)	(0.000)	(0.000)
Time Unemployed	0.982	0.976	0.977
	(0.043)	(0.014)	(0.015)
(Time Unemployed)^2	1.000	1.001	1.001
	(0.123)	(0.053)	(0.054)
Unemployed X Self Duration	1.038	0.966	0.983
	(0.007)	(0.005)	(0.132)
(Unemployed X Self Duration)^2	0.999	0.999	0.999
	(0.024)	(0.226)	(0.036)
N. of Observations	44,447	44,447	44,447
N. of Subjects	5,387	5,387	5,387
N. of Failures	3,589	3,589	3,589
Time at Risk	46,644	46,644	46,644