Abstract: We study the effect of ability and liquidity constraints on entrepreneurship. We develop a three sector Roy model that differentiates between entrepreneurs and other self-employed. The model predicts—and the data confirm—that entrepreneurs are positively selected on highly valued human capital, but other self-employed are negatively selected on those skills; entrepreneurs are positively selected on collateral, but not other self-employed and entrepreneurship is procyclical but self-employment is countercyclical.

Keywords: Entrepreneurship; Human capital; Occupational choice; Corporate finance; Business cycles

JEL Classifications: L26; J24; G32; E32
I. INTRODUCTION

Entrepreneurship plays a central role in influential theories of economic growth and business cycles. For example, Smith (1776), Schumpeter (1911), and Aghion and Howitt (1992) emphasize that entrepreneurs facilitate economic growth by bring new goods, services, and technologies to the economy. Lucas (1978), Baumol (1990), Murphy et al (1991), and Gennaioli et al (2013) stress that the allocation of entrepreneurial talent influences the productivity of firms and the growth rates of economies. On business cycles, Veblen (1904), Fisher (1933), Keynes (1936), Shleifer (1986), Bernanke and Gertler (1989), Caballero and Hammour (1994) and others explain that the response of entrepreneurs to aggregate shocks shapes how those shocks propagate through the economy. Unsurprisingly, therefore, a rich body of research explores selection into entrepreneurship.

Research, however, highlights three puzzling gaps between theory and evidence regarding the human capital, earnings, and liquidity constraints of entrepreneurs. Several leading theories emphasize that (1) entrepreneurs have unique human capital traits—including creativity, analytical skills, risk taking, self-confidence, education, and managerial acumen (Schumpeter 1911, Lucas 1978, Kihlstrom and Laffont 1979, Evans and Jovanovic 1979, Baumol 1990, Murphy et al 1991, and Gennaioli et al 2013); (2) entrepreneurs are highly remunerated for these scarce skills and for the additional risks associated with entrepreneurial endeavors (Lucas 1978 and Kihlstrom and Laffont 1979); and (3) liquidity constraints limit entry into entrepreneurship (Knight 1921, Bernanke and Gertler 1989, Evans and Jovanovic 1989, Kiyotaki and Moore 1997, and Cagetti and De Nardi 2006). Yet, some influential empirical studies find that (1) the typical self-employed person does not have better skills or educational attainment than salaried employees (Fairlie 2002); (2) the typical self-employed person does not earn more than her salaried counterpart (Borjas and Bronars 1989, Evans and Leighton 1989, Hamilton 2000, and Moskowitz and Vissing-Jorgensen 2002); and (3) liquidity constraints restrict only a small proportion of wealthy individuals from becoming self-employed (Hurst and Lusardi 2004).
Researchers have addressed each of these puzzles independently. On the human capital puzzle, while researchers find no evidence for positive selection into self-employment on cognitive skills, several researchers point to selection on noncognitive traits including the degree of risk aversion, break-the-rules mentality, etc. (Fairlie 1999, 2002, Fairlie and Robb 2007a,b, Fairlie and Woodruff 2010, Hartog, van Praag, and van der Sluis 2010, and Nanda and Sørensen 2010). On earnings, some argue that “overly confident” business owners (Bernardo and Welch 2001, De Meza and Southey 1996, and Dawson et al. 2014), the non-pecuniary benefits of self-employment (Hurst and Pugsley 2011), attribution bias (Manso 2016), and underreported income (Hurst, Li, and Pugsley 2014) help explain the “earnings puzzle.” On liquidity constraints—and without necessarily rejecting the Hurst and Lusardi (2004) finding that liquidity constraints bind for few, considerable research shows that (1) entrepreneurial wealth in general shapes entry into self-employment (Evans and Jovanovic 1989, Evans and Leighton 1989, Holtz-Eakin, Joulfaian, and Rosen 1994a, and Fairlie 1999) and (2) housing wealth in particular influences both selection into self-employment (Black, De Meza, and Jeffreys 1996, and Fairlie and Krashinsky 2012, Fort et al. 2013, Corradin and Popov 2015, and Schmalz, Sraer, and Thesmar 2017) and employment in start-up firms (Adelino, Schoar, and Severino 2015).1

Levine and Rubinstein (2017) address the human capital and earnings puzzles by focusing on the incorporated self-employed as a better proxy for entrepreneurship than the aggregate group of self-employed. They document that entrepreneurs possess a unique mixture of cognitive and non-cognitive traits and earn more than the typical salaried worker. Yet, most of the self-employed are not incorporated and these other self-employed have very different traits and earnings. What Levine and Rubinstein (2017) and others do not provide is a theory that (a)

---

1 There is large literature on housing collateral and credit constraints. A growing body of work finds that housing prices influence corporate investments (Gan 2007a and Chaney, Sraer, and Thesmar 2012), bank lending (Gan 2007b), and household access to credit (Gan 2010 and Mian, Rao, and Sufi 2013), though see Kerr, Kerr, and Nanda (2015) for a contrasting view on the importance of housing wealth for entrepreneurship. On collateral, corporate investment, and corporate finance more generally, see, for example, Fazzari, et al. 1988, Hoshi, Kashyap, and Scharfstein (1991), Holtz-Eakin, Joulfaian, and Rosen 1994b, Blanchflower and Oswald (1998), Hubbard and Kashyap (1992), Gertler and Gilchrist (1994), Benmelech and Bergman (2008), and Benmelech, Garmaise, and Moskowitz (2005).
explains why aggregating these different types of self-employed is conceptually inappropriate and will be empirically misleading and (b) simultaneously accounts for the human capital, earnings, and liquidity puzzles. Specifically: (1) Why is there positive selection into entrepreneurship on key human capital traits but negative selection into other forms of self-employment? (2) Why do the incorporated tend to earn more than salaried workers, while the other self-employed earn less than their salaried counterparts? (3) Why do liquidity constraints limit so few from becoming business owners, and do liquidity constraints bind differently for entrepreneurs and the other self-employed?

In this paper, we offer a unified treatment of entrepreneurship and self-employment that addresses these puzzles. In particular, we first develop a theoretical model of how human capital, preferences, and liquidity constraints shape selection into entrepreneurship, other forms of self-employment, and salaried employment. We then, use our framework to analyze empirically the roles of cognitive and noncognitive traits, labor market skills, and credit constraints in shaping selection into the different employment types and use these findings to explain the human capital, earnings and liquidity “puzzles.” Furthermore, we use our framework to address a different and enduring debate about the cyclicity of entrepreneurship, where we allow for pro- and counter-cyclical forces to differentially influence entry into entrepreneurship and other self-employment.

A key starting point in building our model is the growing body of evidence that self-employment is a problematic proxy for entrepreneurship because it fails to distinguish between entrepreneurs and other self-employed individuals. Evans and Leighton (1989), Schoar (2010), Hurst and Pugsley (2011), La Porta and Shleifer (2014), Levine and Rubinstein (2017), and others indicate that some of the self-employed undertake highly-productive ventures that create jobs and introduce new goods and services to the market, i.e., “entrepreneurs.” Most of the self-employed, however, are one-person, low-productivity “other self-employed” individuals, who were unsuccessful salaried workers, perform routine, manuals tasks, and have no ambitions to grow their businesses. Thus, bundling together these two different types of self-employment—
conceptually and empirically—might yield misleading perspectives and inferences about entrepreneurs.

We first develop a three-sector Roy model that distinguishes between entrepreneurs, salaried employees, and other self-employed. Our model differs from Evans and Jovanovic’s (1979) (henceforth EJ) influential model of entrepreneurship in two key respects. While EJ aggregate business owners into one category of self-employment, we distinguish between (i) entrepreneurship—which demands entrepreneurial ability, physical capital, and liquidity—and (ii) other forms of self-employment that demand none (or little) of these and is driven primarily by the non-pecuniary benefits of self-employment, such as being one's own boss, with “… little desire to grow big or to innovate in any observable way” (e.g., Hurst and Pugsley 2011). Our model also differs from EJ in that we relax their assumption that entrepreneurial ability is not valuable in salaried employment. Similar to EJ, our model includes complementarities between human and physical capital that give rise to a Knightian (1921) connection between the entrepreneur and capitalist. That is, the optimal capital stock is increasing in entrepreneurial ability, which means that liquidity constraints are more costly to higher-ability entrepreneurs.

The model yields unique predictions with respect to human capital, earnings, and liquidity constraints. First, entrepreneurs are positively selected on entrepreneurial ability and salaried wages, when entrepreneurial ability is also useful in salaried employment, but the other self-employed are negatively selected on both. Thus, the model highlights the conceptual and potentially empirical problems with aggregating (i) the high-ability entrepreneurs, who earn high-wages when they work as salaried employees, with (ii) the low-ability other self-employed, who earn low-wages when they work as salaried employees. The model’s second prediction is that entrepreneurs are positively selected on collateral and access to capital, but the other self-employed are not. Thus, combining these two types of self-employment may yield an aggregate group in which only a small proportion is liquidity constrained and only a few enter and exit self-employment in response to liquidity shocks.
Our framework also provides distinctive predictions about the cyclicality of entrepreneurship and other self-employment. Several business cycle theories stress that the procyclicality of entrepreneurship amplifies aggregate shocks (e.g., Shleifer 1986, Bernanke and Gertler 1989, Kiyotaki and Moore 1997, Francois and Lloyd-Ellis 2003, and Barlevy 2007). Other models, however, highlight countercyclical forces, emphasizing that the opportunity costs of investment are lower in recessions (Caballero and Hammour 1994) and weak demand for labor in recessions pushes workers temporarily into self-employment (e.g., Kihlstrom and Laffont 1979, and Banerjee and Newman 1993). Empirical assessments of the cyclicality of entrepreneurship—using the aggregate group of self-employed—have not resolved this debate (Evans and Leighton 1989, Parker 2009, Koellinger and Thurik 2012, and Yu, Orazem, and Jolly 2014).

Our model highlights the conceptual problems with using the aggregate group of self-employed to analyze the cyclicality of entrepreneurship. Following the literature, recessions involve both a drop in the demand for salaried workers (reducing the opportunity costs of self-employment) and a tightening of liquidity conditions (increasing the cost of capital). In the model, the drop in demand for salaried workers increases the flow of people into both types of self-employment, i.e., this labor demand effect is countercyclical for both entrepreneurs and other self-employed. The tightening of liquidity conditions, however, impedes people from entering entrepreneurship but has no effect on entry into other self-employment, which demands no (or little) capital. As a result, the model predicts that other self-employment is countercyclical, but entrepreneurship will be procyclical if the liquidity constraint effect is strong enough. Thus, examining the aggregate group of self-employed can hide the distinct cyclical patterns of entrepreneurship and other self-employment.

In turning to the data, we follow Levine and Rubinstein (2017) and use the incorporated as a proxy for “entrepreneurs” and the unincorporated as a proxy for the model’s other group of self-employed individuals. Conceptually, the corporation’s defining legal characteristics—limited liability and a separate legal identity—are most useful for undertaking large, risky
investments that require external financing. Thus, when people establish smaller businesses that do not require much external finance, they will choose the simpler unincorporated legal form; and, when they start larger, risky—more “entrepreneurial”—ventures, they will incorporate. Empirically, Levine and Rubinstein (2017) show that the incorporated and their businesses engage in activities that demand strong nonroutine analytical skills, such as creativity, complex problem-solving, and persuading, motivating, and managing others. In contrast, the unincorporated and their businesses perform activities that demand strong manual skills. To the extent that stronger cognitive skills are more closely aligned with core conceptions of entrepreneurship than strong eye-hand coordination, these results advertise the value of using incorporation is a proxy for entrepreneurship. Furthermore, Levine and Rubinstein (2017) provide evidence that the choice of the business’s legal form reflects the ex ante nature of the underlying endeavor—not selection on the ex post success of the business. That is, very few people start with unincorporated business then incorporate if the endeavor is successful. Although using the incorporated and unincorporated as proxies for entrepreneurs and other self-employed respectively is admittedly crude, there are conceptual and empirical reasons for preferring this demarcation to using the aggregate group of self-employed.

Using the National Longitudinal Survey of Youths 1979, we document that the incorporated and unincorporated are notably different with respect to human capital and starting capital. As teenagers, incorporated business owners have stronger analytical skills, greater self-esteem, and a stronger sense of controlling their futures than those who become unincorporated self-employed. Furthermore, when comparing the salaried wages of people when they were in their 20s (early career salaried wages), those who become incorporated self-employed tend to have higher early career salaried waged than those who either remained salaried workers or switched into unincorporated self-employment. There are also notable differences in starting capital. The typical incorporated business starts with almost ten-times as much capital as the typical unincorporated business, and 21% of the unincorporated report needing no capital to start their businesses.
We begin by evaluating the model’s predictions concerning the differential selection of individuals into incorporated and unincorporated self-employment on entrepreneurial traits, early career salaried wages, and collateral. We first discover that entrepreneurs—as proxied by the incorporated self-employed—are positively selected on a mixture of cognitive and non-cognitive traits and early career salaried wages, while the unincorporated are negatively selected on these same features. Second, we find that entrepreneurs are positively selected on collateral—as measured by home wealth, while the unincorporated are not. Besides being consistent with the model’s predictions, these results offer a resolution of the human capital, earnings, and liquidity puzzles: When researchers combine entrepreneurs with the other self-employed, this aggregates away the unique human capital traits and high earnings of entrepreneurs and obfuscates the connection between entrepreneurship and liquidity constraints.

Next, we exploit natural variation in home equity values across regions and time and the cross-sectional variation in home ownership to identify the impact of collateral, and hence liquidity constraints, on entry into entrepreneurship and other self-employment. Thus, we follow a long literature that stresses that an individual’s housing wealth shapes credit constraints and the ability to start and grow a business (e.g., Chaney, Sraer, and Thesmar 2012, Adelino, Schoar, and Severino 2015, Corradin and Popov 2015, Schmalz, Sraer, and Thesmar 2017). We discover economically large and statistically significant effects of collateral on entry into entrepreneurship, but no effect on entry into other forms of self-employment. Therefore, aggregating entrepreneurs and other self-employed into one homogeneous business category dilutes the estimated impact of liquidity constraints on entrepreneurship. This helps explain the liquidity puzzle.

We then turn to the cyclicality of entrepreneurship and other forms of self-employment. We use cross-year variation in state unemployment rates to assess the cyclicality of entrepreneurship and the other self-employed. Consistent with our model’s predictions, incorporated self-employment is procyclical, unincorporated self-employment is countercyclical, and aggregate self-employment is countercyclical. During periods of high unemployment, entrepreneurship falls, but there is a sharp increase in unincorporated self-employment that
reverses when the economy recovers. This suggests that some people use unincorporated self-employment as a temporary cushion against adverse labor market shocks (e.g., Farber 1999). Since cyclical fluctuations in unincorporated self-employment are larger than those in incorporated self-employment, our findings (a) confirm and account for past findings that aggregate self-employment is countercyclical and (b) uncover the procyclicality of entrepreneurship.

Finally, we extend the model and the empirical analyses to consider risk aversion. The model predicts that the optimal capital stock reflects the interaction between entrepreneurial skills and risk aversion. Thus, effective entrepreneurial ability is a mixture of narrowly defined entrepreneurial skills and noncognitive traits that allow individuals to effectuate those skills. Selection into entrepreneurship, therefore, is determined by the joint distribution of entrepreneurial skills and risk aversion. Thus, some people with exceptional entrepreneurial abilities might choose to work as salaried employees if they are comparatively averse to the risk (and other features associated with entrepreneurship). Consistent with this prediction, we find that among smart, able people—whether measured by cognitive test scores or early career wages—only those who are more likely to engage in risky behaviors, as measured by their tendency to engage in illicit activities as teenagers, select into entrepreneurship.

The remainder of the paper is organized as follows. Section II documents the human capital and liquidity puzzles. Section III presents the theoretical model, and Section IV develops the statistical model, so that we can move from the theory to estimable equations. Section V provides the empirical evaluation of the model’s predictions and Section VI concludes.
II. THE HUMAN CAPITAL AND LIQUIDITY PUZZLES

In this section, we document puzzles concerning the human capital and liquidity constraints of entrepreneurs. We first show that salaried employees and the self-employed have similar human capital traits despite an abundance of theoretical models emphasizing the distinct features of entrepreneurs. Second, we show that most businesses start with less than $3,500 of capital. This is consistent with the findings in Hurst and Lusardi (2004), who stress that liquidity constraints bind for very few. To illustrate these puzzles—and foreshadow our strategy for resolving them, we use data from the NLSY79. We do not document the earnings puzzle here since many researchers, e.g., Hamilton (2000), show that the median self-employed and salaried workers earn about the same per hour despite models emphasizing the unique talents and risk-taking attributes of successful entrepreneurs.

II.A. Data

The NLSY79 is a representative survey of 12,686 individuals who were 15-22 years old when they were first surveyed in 1979. Individuals were surveyed annually through 1994 and biennially since then. Thus, we use year $t-2$ when referring to a lagged value. We examine individuals who are 30 years of age or older for whom the NLSY79 has information on assets, standard demographic information, and the human capital traits described below.

The NLSY79 reports information on human capital. It provides basic demographics, such as age, gender, race, and state of residence. It provides information on educational attainment, including the number of years of education and whether the person graduated from college.

The NLSY79 also contains measures of cognitive ability, illicit activities, and personality traits. From the 1980 survey, AFQT (Armed Forces Qualifications Test) measures the aptitude and trainability of each individual and is often used as an indicator of cognitive skills. The AFQT indicates the individual’s percentile within the entire sample and has a median of 50. Furthermore, we construct the index $Illicit$ that measures the aggressive, risk-taking, disruptive, “break-the-rules” behaviors of individuals before they reach prime working-age. $Illicit$ is based on 20 survey questions from the 1980 NLSY79 that cover actions associated with damaging

---

2 Appendix Table I provides detailed variable definitions and sources.
property, fighting, shoplifting, robbery, assault, drug use and dealing, etc., and whether the individual was stopped by the police, charged with an illegal activity, or convicted of non-minor traffic violations. We construct this index to have a mean of zero and standard deviation of one. In addition, we construct Smart & Illicit, which equals one for an individual if (a) AFQT is 50 or above and (b) Illicit is zero or above. Otherwise, Smart & Illicit equals zero. With respect to personality traits, the Self-Esteem index measures the degree of approval or disapproval of one’s self and is based on ten questions in the 1980 survey. Locus of Control is from the 1979 survey and measures the degree to which individuals believe they have internal control of their lives through self-determination relative to the degree that external factors, such as chance, fate, and luck, shape their lives. Smaller values indicate a greater sense of self-determination. Both Rosenberg Self-esteem and Locus of control are standardized across all individuals in the survey, so that each has a mean of zero and standard deviation of one.

The NLSY79 also contains information on wealth, earnings, and the amounts used to start businesses. We compute Home Wealth as the market value of the individual’s home minus any mortgages on it and Wealth as the value of all assets minus all liabilities. To compute real earnings, the NLSY79 provides nominal earnings, and we use the Consumer Price Index to convert these values into 2010 dollars. Furthermore, we construct Wages (25-29), which equals an individual’s average real log hourly earnings as a salaried employee during the ages of 25 through 29 if the person is 31 years of age or older and equals the individual’s average real log hourly earnings as a salaried employee in t-2 if the person is between the ages of 27 and 30. When people are less than 27 year old, we set Wages (25-29) equal to missing. Wages (25-29) is available for almost all individuals, since people typically start their working lives as salaried workers. Starting with the 2010 survey, the NLSY79 began asking businesses about the amount of capital used to start the business (Starting Capital) and the number of employees (Employees).

With respect to employment types, the NLSY79 classifies all workers in each year as either salaried or self-employed, and among the self-employed, indicates whether individuals are incorporated or unincorporated. Specifically, individuals are asked about the employment class for their main job: “Were you employed by a government, by a private company, a nonprofit organization, or were you self-employed (or working in a family business)?” Those responding
that they are self-employed are further asked, “Is this business incorporated?” While incorporation offers the benefits of limited liability and a separate legal identity, there are direct costs of incorporation, such as annual fees and the preparation of more elaborate financial statements, and indirect costs associated with the separation of ownership and control.

We use the incorporated as a proxy for entrepreneurs and the unincorporated as a proxy for the other self-employed in our model. Levine and Rubinstein (2017) show that the incorporated and their businesses engage in activities that demand a relatively high degree of creativity, complex problem-solving, and communication skills, including the ability to persuade, motivate, and manage others. In contrast, the unincorporated perform activities that require relatively low levels of these analytical skills but instead require strong manual skills. Under the assumption that stronger cognitive skills are more closely aligned with core conceptions of entrepreneurship than manual dexterity, these observations motivate our use of incorporation as a better proxy for entrepreneurship than aggregate self-employment.

II.B. Patterns: Human capital

Table I provides summary statistics on individuals and their businesses. Focusing on those who work full-time, full year, the table differentiates individuals by whether they are salaried employees (Employed) or self-employed. For the self-employed, the table provides summary statistics on all self-employed (Total) and also by the legal form of the business (Unincorporated or Incorporated). The data are from the business ownership part of the 2010 and 2012 NLSY79 surveys.3 For the business ownership part of the surveys, the observation is at the person-business level. Specifically, individuals are classified as business owners based on the 2010-2012 waves. Individuals who are not business owners enter the sample only once. Individuals who are business owners have an entry per business reported. If a person reports one business – she enters once. If a person reports two businesses, she enters twice. Accordingly, data on starting capital and the legal form of the business are per entry.

---

3 There are some differences between the responses that individuals give regarding employment type in the business ownership and employment parts of the NLSY79. In Table 1, we classify an individual as incorporated or unincorporated only if the individual provides consistent responses in both parts of the survey. The results, however, are very similar if we classify employment type based either the business ownership or employment part.
Table I shows that the human capital traits of employees and the self-employed are similar. On average, employees and the self-employed in 2010 and 2012 have virtually the same (a) number of years of education (13.8 v. 13.7), (b) proportion of college graduates (29% v. 28%), (c) salaried earnings when they were 25-29 years old (2.35 for employees and 2.39 for the self-employed). We also compare measures of the cognitive abilities and personality traits of individuals before they entered the prime age workforce. We find that the differences between employees and the self-employed are small, though the self-employed have slightly higher AFQT scores and self-esteem values, and slightly lower values of the Locus of control indicator. For example, there is only a 2.2 percentile point difference in average AFQT scores between employees (49.2) and the self-employed (51.4). Thus, although influential models of entrepreneurship emphasize the unique human capital of entrepreneurs, the self-employed and salaried employees have remarkably similar attributes.

Table I also hints at an explanation of this human capital puzzle: There are two distinct types of self-employed, those who tend to engage in entrepreneurial activities (Incorporated) and those who do not (Unincorporated). These two types of self-employed have very distinct human capital characteristics. The incorporated have, on average, more years of education, a much greater likelihood of graduating from college, and earn much more per hour than both the unincorporated self-employed and salaried employees. In contrast, the unincorporated have lower levels of each of these human capital indicators than salaried employees and incorporated business owners. There are also large differences in AFQT, Self-esteem, and Locus of control. For example, the incorporated have AFQT scores that are, on average, 11.5 percentile points greater than the unincorporated and 10.6 percentile point greater than salaried workers, while the unincorporated have the lowest AFQT scores across employment types. Thus, while salaried employees have similar human capital to the aggregate group of self-employed, entrepreneurs tend to have much greater human capital than salaried workers, while other self-employed have much less. Aggregation may account for the human capital puzzle.
II.C. Patterns: Starting capital

Table I also documents that the median self-employed individual (a) starts the business with less than $3,500 and (b) has no employees. This is consistent with the findings of Evans and Leighton (1989), Hurst and Lusardi (2004), Hurst and Pugsley (2011), and Levine and Rubinstein (2017): most businesses are one-person, retail operations that provide routine, manual services, such as landscaping, house cleaning, handyman services, etc. Indeed 17% of the self-employed indicate that no capital was needed to start their businesses. These observations on starting capital further motivate the question raised by the findings in Hurst and Lusardi (2004): Do liquidity constraints represent a high entry barrier for many potential entrepreneurs?

The notable differences between entrepreneurs and other self-employed may also account for this liquidity puzzle. Table I indicates that the median starting capital for an unincorporated business is about $2,000, but it is almost $20,000 for incorporated businesses. While 21% of individuals report needing no capital to start an unincorporated business, only 5% of incorporated business owners respond similarly. Also note, that the average incorporated business has more than ten-times the number of employees as an average unincorporated business. There are also pronounced differences in wealth. The total wealth of the unincorporated self-employed is, on average, about $70,000, of which $19,500 is home wealth. In contrast, the overall wealth of incorporated business owners is almost $160,000, of which $32,000 is home wealth. These differences in collateral and starting capital suggest that aggregating the incorporated and unincorporated self-employed might yield misleading information on the degree to which liquidity constraints limit entry into entrepreneurship.

Table I documents that contrary to influential theories, salaried employees and self-employed individuals have similar human capital characteristics and most businesses start with little or no capital. The data also suggest a strategy for resolving these puzzles: There are material differences between incorporated and unincorporated self-employed and their businesses. Thus, we now develop a three-sector Roy model to explore the selection of
III. A MODEL OF SELECTION INTO ENTREPRENEURSHIP AND SELF-EMPLOYMENT

III.A. Framework

Each individual chooses one of three employment types: Salaried employment (S), entrepreneurship (E), and other self-employment (U). Individual \( i \) then receives income \( I_{Ji} \) from working in employment type \( J \), where \( J \) is \( S, E, \) or \( U \). Individuals sort into employment types to maximize utility, where the utility of individual \( i \) in employment type \( J \) is a function of income and non-pecuniary benefits \( \delta_{Ji} \):

\[
V_{Ji} = I_{Ji} * e^{\delta_{Ji}}.
\]

The non-pecuniary benefits of self-employment could, for example, reflect preferences to be one’s “own boss,” as emphasized by Hurst and Pugsley (2011). Non-pecuniary benefits are defined relative to salaried employment, so that \( \delta_{Si} = 0 \). We first derive the model with risk neutral individuals, as in Evans and Jovanovic (1979), henceforth EJ, and then extend the model to allow for risk aversion.

Individuals are endowed with human capital, consisting of (1) entrepreneurial ability \( \theta_i \) and (2) other employment specific skills \( \varepsilon_{Ji} \) that are uncorrelated with entrepreneurial ability. Without loss of generality, we assume that \( \theta_i > 0 \), \( \varepsilon_{Ei} = 0 \), and \( \varepsilon_{J} > 0 \), for \( J = S \) or \( U \).

Human capital skills are not equally productive across employment types. Specifically, the effective human capital of individual \( i \) in employment type \( J \) \( (H_{Ji}) \) is

\[
H_{Ji} = \theta_i^{\rho_J} * e^{\varepsilon_{Ji}},
\]

where the effective human capital of entrepreneurial ability in employment type \( J \) is represented by \( \rho_J \), so that it is natural to set \( \rho_E = 1 \). While EJ assume that \( \rho_S = 0 \), we relax this assumption and allow abilities that are useful for entrepreneurship to also be productive in salaried employment. Thus, we assume that \( \rho_U < \rho_S \leq 1 \), and without further loss of generality, set
\( \rho_U = 0 \). We evaluate these assumptions empirically below.\(^4\) Thus, the effective human capital of individual \( i \) in salaried employment is increasing in (a) the person’s job-specific skills in salaried work (\( \epsilon_{S_i} \)), (b) the persons entrepreneurial ability (\( \theta_i \)), and (c) the degree to which entrepreneurial ability is productive in salaried employment (\( \rho_S \)).

An individual choosing employment type \( U \) or \( S \) earns
\[
I_{ji} = H_{ji},
\]
which expressed as log earnings is
\[
\ln I_{ji} = \rho_j \ln \theta_i + \epsilon_{ji}.
\]

Individuals engaged in entrepreneurship combine entrepreneurial ability and physical capital (\( K \)) to produce output (\( Y \)) using a similar production function as in EJ:
\[
Y_i = H_{EI} K_i^\alpha v_i = \theta_i K_i^\alpha v_i,
\]
where \( 0 < \alpha < 1 \) and \( v_i \) is a lognormal disturbance that reflects an independent and identically distributed productivity shock, where \( E[v_i] = 1 \). As in Lucas (1978), Jovanovic (1982), EJ, and many others, entrepreneurs with more entrepreneurial ability have, ceteris paribus, larger average and marginal products of capital at each level of capital.

Net returns from entrepreneurship, i.e., entrepreneurial earnings (\( I_{EI} \)), equal
\[
I_{EI} = \theta_i K_i^\alpha v_i - r_i K_i,
\]
where the price of output is one and the gross cost of capital (\( r_i \))—one plus the interest rate—is heterogeneous across individuals and greater than one. For now, we simply take \( r_i \) as given. In particular, we assume individuals are endowed with exogenously given assets and when these assets are used as collateral to finance \( K \), they reduce the cost of capital. This is a bit different from EJ, where collateral determines how much an individual can borrow. In our model,

\(^4\) This assumption that \( \rho_S > \rho_U \) is consistent with the findings in Levine and Rubinstein (2017). They show that the pecuniary returns to human capital skills that are crucial for entrepreneurial success (e.g., learning aptitude scores) are greater in salaried employment than in other self-employment, suggesting that \( \rho_S > \rho_U \). Given this assumption, setting \( \rho_U = 0 \) is just a simplifying normalization that does not affect the analyses. In the model, if \( \rho_S = 0 \), then there is no unique level of entrepreneurial skill at which individuals are indifferent between \( U \) and \( S \).
collateral influences the cost of capital and endogenously influences the optimal capital stock and hence borrowing.\(^5\)

The \(K_i\) that maximizes expected entrepreneurial earnings \((K_i^*)\), given \(\theta_i\) and \(r_i\), is

\[
K_i^* = (\theta_i \alpha / r_i)^{1/(1-\alpha)},
\]

and the log of expected entrepreneurial earnings at this maximum is therefore:

\[
ln I_{E_i}^* = \rho'_E ln \theta_i + \alpha \rho'_E \ln \left( \frac{\alpha}{r_i} \right) + \ln(1 - \alpha),
\]

where \(\rho'_E = \left( \frac{1}{1-\alpha} \right)\).

Notice three features about entrepreneurial earnings. First, entrepreneurial earnings (and the optimal capital stock) are increasing in entrepreneurial ability \(\theta_i\) and decreasing in the cost of capital \(r_i\). Second, the elasticity of entrepreneurial earnings with respect to entrepreneurial ability is greater than one, i.e., \(\rho'_E > 1\). This reflects the endogeneity of capital to entrepreneurial ability: Higher \(\theta_i\) not only increases the returns to entrepreneurship at each level of capital, it increases the returns to increasing the capital stock. Third, by comparing equations (4) and (8), note that the returns to entrepreneurial ability in entrepreneurship are larger than the returns to entrepreneurial ability in salaried employment even when \(\rho_S = 1\). This arises because of the complementarity between entrepreneurial ability and physical capital.

\(^5\) Under the same assumptions, all of the results hold when using the EJ formulation of liquidity constraints.
III.B. Selection into employment types

Individuals select into employment types $U$, $S$, or $E$ by comparing expected utility levels:

\[
\ln V_{Ul} = \varepsilon_{Ul} + \delta_{Ul}, \tag{10.1}
\]
\[
\ln V_{Sl} = \rho_S \ln \theta_l + \varepsilon_{Sl}, \tag{10.2}
\]
\[
\ln V_{El} = \rho_E' \ln \theta_l + \alpha \rho_E' \ln \left( \frac{a}{r_i} \right) + \ln(1 - \alpha) + \delta_{El}. \tag{10.3}
\]

In comparing the logs of expected utilities across employment types, note the following results on human capital. First, utility in entrepreneurship rises faster in $\theta_l$ than utility rises in either salaried work or other self-employment. This holds even when entrepreneurial abilities are equally productive in salaried work, namely $\rho_S = \rho_E = 1$ because of the complementarity between entrepreneurial ability and physical capital within entrepreneurial endeavors. Second, the log of utility in salaried employment ($\ln V_{Sl}$) reflects both human capital that is specific to salaried employment ($\varepsilon_{Sl}$) and entrepreneurial human capital that is valuable in salaried work ($\rho_S \ln \theta_l$). One implication of these first two results is that, ceteris paribus, increases in $\varepsilon_{Sl}$ boost the relative utility of salaried employment, but increases in $\ln \theta_l$ boost the relative utility of entrepreneurship because $\rho_E' > 1 \geq \rho_S$. Third, liquidity constraints, in the form of higher cost of capital ($r_i$), reduce proportionally the utility from entrepreneurship, and therefore have larger adverse effects on able entrepreneurs. The overall effect reflects the direct cost of capital and its indirect effect on the optimal level of capital and gross earnings. Fourth, other human capital endowments ($\varepsilon_{JI}$) and preferences ($\delta_{JI}$) directly shape the relative utility of different employment types.

We now derive the cutoff levels of entrepreneurial ability that lead individuals to select into self-employment, salaried employment, or entrepreneurship. We derive these cutoff levels of $\ln \theta_l$ as functions of the cost of capital ($r_i$), non-entrepreneurial human capital skills ($\varepsilon_{Ul}$ and $\varepsilon_{Sl}$), preferences ($\delta_{JI}$), and the degree to which entrepreneurial ability is remunerated in salaried employment ($\rho_S$) and entrepreneurship ($\rho_E'$).
First, define $\ln \theta_{sI}$ as the level of entrepreneurial ability ($\ln \theta_i$), such that the individual is indifferent between other self-employment and salaried work. Below $\ln \theta_{sI}$, the individual prefers other self-employment to salaried work; and above $\ln \theta_{sI}$, the individual prefers salaried work to other self-employment. Setting $\ln V_{ul} = \ln V_{sI}$, and solving for $\ln \theta_{sI}$ yields:

$$\ln \theta_{sI} = \frac{\delta_{ul} + (\varepsilon_{ul} - \varepsilon_{sI})}{\rho_s}. \quad (11.1)$$

Individuals with stronger preferences for self-employment, $\delta_{ul}$, (e.g., people who like being their own bosses) will have higher $\ln \theta_{sI}$ cutoff values than otherwise similar individuals. Furthermore, in economies where entrepreneurial ability is more highly remunerated in salaried employment (higher $\rho_s$), $\ln \theta_{sI}$ will be correspondingly lower, because it takes less entrepreneurial ability to generate the earnings level in salaried employment that makes the individual indifferent between salaried work and other self-employment. Finally, note that if there are insufficient pecuniary ($\varepsilon_{ul}$) and nonpecuniary returns ($\delta_{ul}$) to other self-employment ($U$) (or skills as a salaried worker are sufficiently high ($\varepsilon_{sI}$)), then $\ln \theta_{sI} \leq 0$ and individuals will not sort into $U$.

Next, define $\ln \theta_{Ei}$ as the level of entrepreneurial ability ($\ln \theta_i$), such that the individual is indifferent between salaried work and entrepreneurship. Below $\ln \theta_{Ei}$, the individual prefers salaried work to entrepreneurship; and above $\ln \theta_{sI}$, the individual prefers entrepreneurship to salaried work. Setting $\ln V_{ei} = \ln V_{sI}$, and solving for $\ln \theta_{Ei}$ yields:

$$\ln \theta_{Ei} = -\frac{\alpha \rho_{E} \ln(\alpha/r_i) + \ln(1-\alpha) + (\delta_{Ei} - \varepsilon_{sI})}{\rho_{E} - \rho_s}. \quad (11.2)$$

Equation (11.2) indicates that individuals facing a higher cost of capital are less likely to become entrepreneurs, implying that individuals with higher $\eta_1$ require more entrepreneurial ability to enter entrepreneurship than similar individuals with lower capital costs. Equation (11.2) also indicates that individuals with greater salaried-specific human capital (larger $\varepsilon_{sI}$) or those receiving less utility from entrepreneurship (smaller $\delta_{Ei}$) will require greater entrepreneurial skills to prefer entrepreneurship over salaried employment.
There are two scenarios. First, when \( \ln \theta_{El} > \ln \theta_{Si} \), the benchmark case, were all employment types are relevant, namely, there are levels of entrepreneurial abilities under which each employment type is optimal. Second, when \( \ln \theta_{El} < \ln \theta_{Si} \), the non-benchmark case, were paid employment is never optimal.

Consider the benchmark case where the cost of capital \((r_i)\), non-entrepreneurial human capital skills \((\varepsilon_{ji})\), preferences \((\delta_{ji})\), and the production function parameter \((\alpha)\) are such that individuals might select into each of the three employment types—self-employment, salaried work, and entrepreneurship—for different values of entrepreneurial ability \( \ln \theta_{El} \). That is, the benchmark involves values of \( r_i, \varepsilon_{ji}, \delta_{ji}, \) and \( \alpha \), such that \( \ln \theta_{El} > \ln \theta_{Si} \), as discussed above.

For the benchmark case, Figure I illustrates the relationship between the log of the expected utility in each employment type and \( \ln \theta \). The horizontal line represents the log of expected utility of other self-employment \((\ln V_{Ui})\) and equals \( \varepsilon_{Ui} + \delta_{Ui} \). The upward sloping line with squares is the log of expected utility of salaried employment \((\ln V_{Si})\), where the slope is \( \rho_S \). \( \ln V_{Si} \) intersects \( \ln V_{Ui} \) at the first cutoff level of entrepreneurial ability: \( \ln \theta_{Si} \). The upward sloping line with circles is the log of the expected utility of entrepreneurship \((\ln V_{El})\), where the slope is \( \rho_E \), and where \( \ln V_{El} \) intersects \( \ln V_{Si} \) at the second cutoff level: \( \ln \theta_{El} \). Except where explicitly noted, we focus on this benchmark case.

Figures I and II illustrate key features of the model under these benchmark conditions. Figure I shows how human capital shapes selection into different employment types. On human capital, entrepreneurs are positively selected on entrepreneurial ability, but the other self-employed are negatively selected on \( \ln \theta_i \). On liquidity constraints, Figure II indicates that \( r_i \) shapes entry into entrepreneurship, but not into self-employment. In particular, increases in \( r_i \) shift downward the intercept of the line for the log utility of entrepreneurship, constraining entry into entrepreneurship. Changes in \( r_i \), however, do not alter the intercepts or slopes of the other lines and therefore liquidity constraints do not affect entry into other self-employment.

Figure I shows that entrepreneurs expect to earn more than salaried workers when the non-pecuniary benefits from entrepreneurship are low (for example \( \delta_{Ui} \leq 0 \)), but this is not the
case for the other self-employed, especially given the non-pecuniary benefits from self-
employment documented by Hurst and Pugsley (2011). Finally, note that the model illustrates
the problems with aggregating the entrepreneurs (E) and the other self-employed (U). The typical
self-employed individual in this aggregate group is not selected on entrepreneurial traits; does
not earn more than the typical salaried worker; and does not face binding liquidity constraints,
since entrepreneurs are a small proportion the aggregate group of self-employed individuals.

III.C. Testable implications and discussion

The model yields testable implications with respect to the impact of human capital,
liquidity, and the business cycle on entrepreneurship and other self-employment. In this
subsection, we highlight three testable implications for the benchmark case, i.e., for the non-
degenerative cases, in which individuals can feasibly sort into each of the three employment
types depending on their entrepreneurial ability.

The first two testable implications relate to section on entrepreneurial ability and salaried
wages. First, there is negative selection on entrepreneurial ability into other self-employment; but
positive selection into entrepreneurship. Second, there is negative selection on salaried wages
into other self-employment; but potentially positive selection on salaried wages into
entrepreneurship when entrepreneurial ability is highly productive in salaried work (e.g., if
$\rho_S = 1$).

Figure I illustrates both of these implications. Individuals with entrepreneurial abilities
above $ln\theta_{ei}$ have better salaried job opportunities and even better entrepreneurial opportunities
than otherwise similar people with lower entrepreneurial abilities. The opposite is true of people
who sort into other self-employment, i.e., the U-employment type. Ceteris paribus, it is people
with lower entrepreneurial abilities and hence people with comparatively low-paying salaried
options, who choose other self-employment. As for selection into entrepreneurship on salaried
wages, this depends on the importance of entrepreneurial abilities in paid-employment. For
example, in an economy where only one skill determines people’s productivity in both salaried
employment and entrepreneurship, people with the best salaried job opportunities become entrepreneurs. As illustrated in Figure II, for $\rho_S = 1$ and $\varepsilon_s = 0$, all other things equal, the most productive salaried workers become entrepreneurs. The positive selection into entrepreneurship reflects the complementarity between ability and capital in entrepreneurship and the lack of perfect adjustment of capital to human-capital in paid employment.

Yet, when entrepreneurial abilities are not very useful in paid-employment (low $\rho_S$), then there can be negative selection on salaried wages. For example, when $\rho_S = 0$ and wage differences reflect only $\varepsilon_s$, as in the EJ model, then there is negative selection on wages into entrepreneurship. Thus, our model allows for positive and negative selection into entrepreneurship on salaried wages depending on the sources of variation in salaried earnings.

The third testable implication is that entrepreneurs are negatively selected on the cost of capital, but the other self-employed are not. As illustrated in Figure II, an increase in the cost of capital implies a parallel drop in the line representing the log utility of entrepreneurship ($lnV_{EL}$). This implies a higher entrepreneurial ability threshold with respect to selection into entrepreneurship but has no effect on selection into other forms of self-employment.

A fourth testable implication involves the cyclicality of entrepreneurship and other self-employment. In the context of our model, we characterize the manifestation of aggregate fluctuations as changes in both the demand for salaried employees and the severity of liquidity constraints. For example, we characterize recessions as a simultaneous reduction in labor demand and a tightening of credit constraints. This is illustrated in Figure III. The tightening of liquidity constraints involves a parallel fall in the log utility of entrepreneurship line ($lnV_{EL}$). As shown, this tightening reduces selection into entrepreneurship but has no effect on entry into other self-employment, i.e., the liquidity effect exerts a procyclical influence on entrepreneurship, but not on other self-employment. With respect to labor demand, a reduction in the demand for salaried employees implies a parallel drop in the line depicting the log utility of salaried employment ($lnV_{Sl}$). The labor demand effect is countercyclical for both types of self-employment. Thus, the model yields (a) an ambiguous prediction about the cyclicality of
entrepreneurship but (b) an unambiguous prediction that other self-employment is countercyclical. It is worth noticing that the aggregate proportion of self-employment might be countercyclical even if entrepreneurship is procyclical.

These implications of the model are unique. Other models of entrepreneurship do not distinguish between entrepreneurs and other self-employed individuals. Therefore, they do not derive predictions regarding the contrasting selection of individuals into entrepreneurship and other self-employment. Our model explains why aggregating these two groups and calling the combined group “entrepreneurs” can lead to misleading perspectives on entrepreneurship. In addition, our model’s prediction that entrepreneurs might be positively selected on salaried wages is very different from EJ, where the less able salaried workers select into entrepreneurship.

III. D. Extension: Risk aversion

We now generalize the utility function to allow for risk-averse individuals. In particular, consider the constant relative risk aversion utility function:

\[ V_{JI} = -\exp \left\{ -\tau_i I_{JI} * e^{\delta_{JI}} \right\}, \]

(1')

where \( \tau_i \) represents individual i’s risk aversion, as defined by \( -V_i'' / V_i' \). Equation (1’) converges to the risk neutral utility function defined by equation (1) as \( \tau_i \to 0 \). Furthermore, we slightly modify the specification of the shock to productivity, so that

\[ Y_i = \theta_i K_i^\alpha (1 + \nu_i'), \]

(5’)

where \( \nu_i' \) is a zero mean, normally distributed shock to productivity. Assuming that the variance of output is \( \sigma_{\eta_i}^2 = \sigma^2 \theta_i K^\alpha \), so that the variance of aggregate output does not change if a firm is split into two or more firms, then the variance of \( \nu_i' \) equals \( \sigma_{\eta_i}^2 / \theta_i K^\alpha \).

Thus, expected utility in entrepreneurship is:

\[ E\{V_{EI}\} = -\exp\{ -\tau_i [\theta_i K_i^\alpha - r_i K_i - \theta_i K_i^\alpha (\sigma^2 / 2)] \}, \]
where, for simplicity, we have set $\delta_{El} = 0$ (rather than $\delta_{Sl}$). Exploiting the observation that the certainty equivalent earnings from entrepreneurship is $I'_{El} = \theta_i K_i^a (1 - \tau_i (\sigma^2 / 2)) - \tau_i K_i$, the optimal capital stock for entrepreneur $i$ is:

$$K_i^* = \left( \frac{\theta_i \gamma_i \alpha}{\rho_E} \right)^{1/(1 - \alpha)}, \quad (7')$$

where $\gamma_i = (1 - \tau_i (\sigma^2 / 2))$, so that $\gamma_i$ is increasing in risk tolerance and decreasing with risk.

The log of the certainty equivalent earnings from entrepreneurship, $ln\{I'_{El}\}$, evaluated at the optimal capital stock is then given by:

$$ln\{I'_{El}\} = \rho_E \ln[\theta_i \gamma_i] + \alpha \rho_E Ln \left( \frac{\alpha}{\gamma_i} \right) + Ln(1 - \alpha). \quad (10.2')$$

Furthermore, since there is no income uncertainty associated with salaried employment or (non-entrepreneurial) self-employment, expected utilities from these employment types are the same as specified above.

Allowing for risk aversion, therefore, yields the following insights. First, the core predictions from the benchmark, risk-neutral specification hold: (1) entrepreneurs are positively selected on entrepreneurial ability ($ln\theta_i$), but other self-employed are negatively selected on entrepreneurial ability, (2) entrepreneurs are positively selected on salaried wages when productivity in salaried employment are highly correlated with entrepreneurial abilities (i.e., when $\rho_S$ is sufficiently large), but other self-employed individuals are negatively selected on salaried wages, and (3) entrepreneurs are negatively selected on the cost of capital but other self-employed individuals are not.

Second, risk aversion reduces the optimal capital stock—and hence the efficiency of entrepreneurial activity. In particular, the optimal capital stock reflects the interaction between ability ($\theta_i$), and attitudes toward risk ($\tau_i$), weighted by risk ($\sigma^2$), i.e., $\theta_i \gamma_i$. This interaction suggests that “effective” entrepreneurial human capital is a mixture of narrowly defined entrepreneurial ability and personality traits that allow individuals to effectuate those skills. Self-selection into entrepreneurship depends on the joint distribution of entrepreneurial abilities ($\theta$) and attitudes toward risk ($\tau$). Thus, the most successful entrepreneurs might not be those with the
most entrepreneurial ability, e.g., if risk tolerance ($\tau$) and entrepreneurial ability ($\theta$) are negatively correlated.\(^6\) This is akin to the combination of “smart and illicit” traits emphasized by Levine and Rubinstein (2017), where illicit captures attitudes toward breaking from the norm, undertaking novel endeavors, and investing in risky ventures.\(^7\)

IV. STATISTICAL MODEL

As discussed above, it is puzzling that existing theoretical models emphasize the crucial roles of both human capital and liquidity constraints in shaping selection into entrepreneurship, but existing empirical research finds that (1) the aggregate group of self-employed has very similar human capital traits and earnings to their salaried counterparts and (2) it takes little capital to start most U.S. businesses. Our model suggests that these findings might reflect the aggregation of entrepreneurs and other self-employed into one category when selection into these two employment types differs systematically on human capital traits, labor market skills, and liquidity constraints.

In this section, we take the theoretical model from section III and derive estimable equations that will allow us to identify statistically and quantify empirically the roles of human capital traits, salaried employment opportunities, and liquidity in shaping selection across employment types and entry into entrepreneurship. In moving from the model toward an estimable equation we need proxies for entrepreneurial traits, salaried employment opportunities, and liquidity constraints. First, with respect entrepreneurial traits, we follow Levine and Rubinstein (2017) and use the interaction between cognitive and non-cognitive traits (“smart and

\(^6\) Therefore, selection on entrepreneurial ability might vary across industries if $\sigma^2$ differs across industries.

\(^7\) This perspective on effective entrepreneurial ability suggests additional research avenues. For example, it might help explain the gender gap in entrepreneurship if women tend to be more risk averse than men as some research documents (e.g., Halevy 2007 and Borghan et al. 2009). As a second example, the model suggests that when risk is smaller, selection into entrepreneurship will be determined more by pure entrepreneurial ability. This might help explain cross industry (and cross country) differences in the human-capital qualities of entrepreneurs and the performance of their businesses.
illicit”) measured early in life, which they show shape selection into entrepreneurship and success as an entrepreneur. This is consistent with our risk-aversion model that highlights the non-separability of entrepreneurial abilities and non-cognitive skills in shaping selection into entrepreneurship. Second, we exploit the observation that almost all individuals work as paid-employees before becoming business owners and use these early career wages to proxy for salaried employment opportunities later in life. Third, on liquidity constraints, we note that home equity is frequently used as collateral to obtain loans. Thus, we use home equity as a proxy for collateral and hence the cost of capital facing an individual.

Assuming that the cost of capital for individual $i$ diminishes with the person’s collateral ($C_{it}$) in the following form $\alpha / r_{it} = \exp(\kappa C_{it})$, where $\kappa$ is a positive constant, and letting $SIL_i$ represent the interaction between cognitive ability (“smart”) and non-cognitive attitudes (“illicit”) of individual $i$, then the probability that individual $i$ prefers entrepreneurship or other self-employment to salaried work is:

$$P(V_J > V_{Sl}) = P(\beta_J W_J + \beta_{JSIL} SIL_i + \beta_J C_{it} + \beta_J X_{it} > \eta_{Sl} - \eta_{Jit}).$$  \hspace{1cm} (11)

$W_J$ represents person $i$'s salaried employment opportunities. As discussed further below, we proxy for $W_J$ using person $i$’s early career wages, i.e., wages between the ages of 25 and 29. $C_{it}$ represents the collateral of person $i$ at time $t$, which, as described below, is proxied for using the equity value of the person’s home. $X_{it}$ is a vector of observable characteristics, including demographics, schooling and early measures of cognitive and non-cognitive traits that might influence employment choices. The error term ($\eta_{Jit}$) combines person-specific shocks to productivity in employment type $J$ in period $t$ and taste, $\eta_{Jit} = \epsilon_{Jit} + \delta_{Jit}$. This directly motivates an estimable multinomial logit regression:

$$\ln(P_{Jit}/P_{Sit}) = \beta_J W_J + \beta_{JSIL} SIL_i + \beta_J C_{it} + \beta_J X_{it},$$ \hspace{1cm} (12)
where the dependent variable is the log-odds ratio of the probability of person $i$ being an entrepreneur ($J=E$) or other self-employed ($J=U$), that is ($P_{Jit}$), rather than a salaried worker ($P_{Slt}$) at time $t$.

There are three reduced form parameters of interest: selection on (i) “smart and illicit” ($\beta_{JSJ}$), (ii) salaried wages ($\beta_{JW}$) and (iii) collateral ($\beta_{JC}$). With respect to cognitive and non-cognitive traits, the model predicts that $\beta_{USIL} < 0$ and $\beta_{ESIL} > 0$. That is, the model predicts that smart and illicit traits are negatively associated with entry into other self-employment (the U-employment type) and positively associated with entry into entrepreneurship. With respect to wages, the model predicts that $\beta_{UW} < 0$: increases in wages increase the utility of salaried employment relative to other self-employment. The model, however, generates ambiguous predictions with respect to $\beta_{EW}$. To the extent that wages are higher because the individual has higher salaried-specific skills, then $\beta_{EW} < 0$: wages rise but entrepreneurial earnings do not. However, when productivity as a salaried worker is sufficiently positively associated with entrepreneurial ability (high $\rho_S$), then the model predicts positive selection into entrepreneurship on wages ($\beta_{EW} > 0$). Indeed, Appendix Table II provides empirical results suggesting that productivity as a salaried worker and entrepreneurial ability are highly correlated.\(^8\) With respect to collateral, the model predicts $\beta_{UC} = 0$ and $\beta_{EC} > 0$. That is, collateral does not shape directly barriers to becoming a salaried worker or U-self-employment type, but collateral lowers the costs of becoming an entrepreneur.

---

\(^8\) Appendix Table II presents regressions of log hourly earnings of individual $i$ in year $t$ on the average log hourly salaried wages of the individual between the ages of 25 and 29 (Wages (25-29)) while conditioning on Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Rosenberg self-esteem, Rotter Locus of Control), as well as race, gender, year, and state fixed effects. We run this regression by employment type in year $t$. We find that early career salary wages are positively related to later entrepreneurial earnings and future salaried earnings but not to the earnings in unincorporated self-employment.
V. SECTION IN ENTREPRENEURSHIP AND OTHER SELF-EMPLOYMENT

In this section, we empirically evaluate the effect of human-traits, early career wages and collateral on selection into entrepreneurship, salaried employment, and other self-employment.

V.A. Section on wages and home wealth

We begin by examining differential selection into incorporated self-employment (entrepreneurship) and unincorporated self-employment (other self-employment) on cognitive and non-cognitive skills, early careers wages, and collateral. Based on equation (12), we focus on estimating the multinomial logit regression:

\[
\ln \left( \frac{P_{JE}}{P_{JS}} \right) = \beta_{JW}W_t + \beta_{JS}SIL_t + \beta_{JC}C_t + \beta_{JX}X_t, \tag{12'}
\]

where the dependent variable is the log-odds ratio of being incorporated \((J=E)\) or unincorporated \((J=U)\) rather than a salaried worker and the other terms are defined above. In Table II, we provide the multinomial logit results on unincorporated and incorporated self-employment (columns 2-3), where we do not report results on other self-employment categories such as unpaid family and nonprofit businesses. In column 1, we also provide the results from a logit regression in which the dependent variable is an indicator variable that equals one if the individual is self-employed (either incorporated or unincorporated) in year \(t\) and zero otherwise.

The key explanatory variables are as follows. For potential salaried wages \((W_t)\), we use \(Wages\ (25-29)\), which equals log hourly salaried earnings when the individual was 25-29 years. For \(SIL_t\), we use \(Smart & Illicit\), which equals one if an individual has both above the NLSY79 sample median of \(AFQT\) and \(Illicit\) and zero otherwise. For \(C_t\), we use \(Home Wealth(t-2)\), which equals the market value of the individual’s home (if any) minus mortgages on the house divided by $100,000 two year before period \(t\). For \(X\), we use the following controls that are not reported in the tables: Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories),\(^9\) measures of cognitive and non-cognitive traits

---

\(^9\) The six educational attainment categories are: (i) high school dropouts: less than 12 years of schooling (ii) GED degree (iii) high school graduates: 12 years of schooling (iv) had some college education: 13-15 years of schooling
(AFQT, Self-esteem, Locus of Control), as well as gender-year, race-year, and state fixed effects. Since the data on home wealth begins in sample year 1985 and we restrict the sample to individuals with data on home wealth in \( t-4 \), the sample starts in 1989. The sample also excludes individuals who were self-employed in either \( t-2 \) or \( t-4 \). The table provides heteroskedasticity robust standard errors clustered at the individual level.

Consistent with the model’s predictions, the results reported in Table II indicate positive selection into incorporated self-employment on early career wages and negative selection into unincorporated self-employment on those wages. \( Wages (25-29) \) enters positively and significantly when examining selection into incorporated self-employment but negatively and significantly when assessing entry into unincorporated self-employment. The economic magnitudes are substantial. Using the estimates from the multinomial logit regressions, consider two people: a high early-career wage earner, where \( Wage (25-29) \) is 25% above the sample median and a low early-career wage earner, where \( Wage (25-29) \) is 25% below the sample median. The coefficient estimates suggest that the odds of the high early-career wage earner switching from salaried work into incorporated business ownership next period are approximately 17% greater than the low early-career wage worker \((1.17=\exp(0.5*0.3139))\). Similarly, the estimated coefficients indicate that the odds of the low early-career wage earner switching from salaried work into unincorporated self-employment next period are 20% greater than the high early-career wage worker \((1.2=\exp(0.5*0.3713))\). Table II also highlights the pitfalls of using the aggregate group of self-employed business owners. As shown in column (1), there is negative selection into aggregate self-employment on early career wages, which masks the differential selection into entrepreneurship and other self-employment.

Also consistent with the model’s predictions, we find positive selection into incorporated self-employment on collateral, but no link between collateral and entry into unincorporated self-

\[ i \text{ college education: 16 years of schooling} \]
\[ vi \text{ advanced studies: 17+ years of schooling} \]
\[ pwe \text{ equals age minus years of schooling minus six (or zero if this computation is negative). The quartic includes } pwe, pwe^2, pwe^3, \text{ and } pwe^4. \]
employment. That is, \textit{Home Wealth}({\textit{t-2}}) enters positively, significantly, and with an economically large coefficient when examining incorporated self-employment but enters with a small, insignificant coefficient when examining unincorporated self-employment. With respect to the economic size of the estimated coefficients, consider a high-collateral and low-collateral person, where the high-collateral person has $50,000 of additional home wealth in year \textit{t-2} than the low-collateral person. The coefficient estimates suggest that the odds of the high-collateral person switching into incorporated business ownership next period from salaried employment this period) are 6.5\% greater than the low-collateral person (=exp(0.5*(0.1607-0.0344)).

The findings on \textit{Smart & Illicit} are also consistent with the model and the findings in Levine and Rubinstein (2017). Like \textit{Wages (25-29)}, \textit{Smart & Illicit} is positively associated with entry into entrepreneurship but negatively associated with entry to unincorporated self-employment. To the extent that \textit{Smart & Illicit} is an additional proxy for effective entrepreneurial abilities that is imperfectly correlated with \textit{Wages (25-29)}, these results are fully in line with the model’s broad predictions. The combination of strong analytical skills and break-from-the-norm, risk-tolerant preferences is positively associated with expected success and hence entry into entrepreneurship. However, these \textit{Smart & Illicit} traits are not productive, and might even be counterproductive, for undertaking the manual-skills-based self-employment activities associated with unincorporated self-employment.
V.B. Section on wages and home wealth: Individual fixed effects and a falsification test

We next address the concern that omitted time-invariant individual traits drive the results on home wealth. For example, if individuals from rich families have characteristics that facilitate both entry into entrepreneurship and larger home equity stakes, then the Table II results might lead us to conclude inappropriately that collateral shapes entry into entrepreneurship when it is the other characteristics that drive both.

To address this concern, we first include individual fixed effects and provide these results in Table III. In this way, we focus on whether changes in collateral are associated with selection into different employment types. Of course, including individual fixed effects will essentially eliminate Wages (25-29), as it varies little over time. As explained in Section II, there is slight time variation in Wages (25-29) when individuals are between 27 and 30 years old. When conducting these analyses with individual fixed effects, we use OLS, because the multinomial logit regressions did not converge. For comparison purposes, therefore, we also present the key earlier analyses from Table II using a linear probability model. Table III shows that the results are robust to conditioning on individual fixed effects. We continue to find positive selection into entrepreneurship on Home Wealth(t-2) but little relation between selection into unincorporated self-employment and Home Wealth(t-2).

We also provide a falsification test in Table III. Instead of examining selection into employment types in period t on home wealth in period t-2, we examine selection into employment types in period t on future home wealth in period t+2 (i.e., on Home Wealth(t+2)). If Home Wealth(t-2) captures changes in wealth that can be used as collateral to finance entry into entrepreneurship in year t, then the model predicts that Home Wealth(t-2) will be positively associated with entry into incorporated self-employment. We would not, however, expect that a change in future household wealth would influence past entry into entrepreneurship unless Home Wealth(t+2) is capturing something else about the evolving characteristics of the individual. When controlling for individual effects, we find positive selection into entrepreneurship on Home Wealth(t-2) but not on Home Wealth(t+2). The results from this falsification test are
consistent with the view that (a) home wealth is positively related to collateral and (b) collateral is important for entering entrepreneurship.

V.C. The impact of home wealth on entry into self-employment

Although the results reported in Tables II and III indicate strong positive selection into entrepreneurship on wages and collateral and strong negative selection into unincorporated self-employment on wages, the analyses do not identify the impact of collateral on entry into entrepreneurship. In particular, lagged housing wealth, even when including individual effects, might not represent an exogenous source of variation in collateral if other time-varying factors shape both home wealth and entry into self-employment.

In this section, we use a Bartik-type instrumental variable to evaluate the impact of collateral on entry into incorporated and unincorporated self-employment. Building on the work in Hurst and Lusardi (2004), Corradin and Popov (2015), and Schmalz, Sraer, and Thesmar (2017), we use \( H_{H_{e}W_{e}W_{l}W_{h}J_{i-4}} * g(i_{-4}, i_{-1}) \), which equals the net value of the home owned by individual \( i \) in year \( t-4 \) \( (H_{H_{e}W_{e}W_{l}W_{h}J_{i-4}) \) times the growth rate of home prices in the state in which the home is located from year \( t-4 \) to year \( t-1 \) \( (g(i_{-4}, i_{-1}) \). If the individual does not own a home in year \( t-4 \), \( H_{H_{e}W_{e}W_{l}W_{h}J_{i-4}} \) equals zero. Our identifying assumptions are that the value of a person’s home in \( t-4 \) and the growth rate in state housing prices between \( t-4 \) and \( t-1 \) are exogenous to the individual’s decision in year \( t \) about switching into incorporated or unincorporated self-employment.

Our identification strategy has two advantages over previous authors. Since we have panel data, we (a) use individual fixed effects as discussed above and (b) conduct the analyses based on a person’s home wealth in \( t-4 \) and changes in the state’s home values from \( t-4 \) to period \( t-1 \) in examining self-sorting into self-employment in period \( t \).

We estimate the following multinomial logit model and report the results in Table IV:

\[
\ln(P_{j_{it}}/P_{s_{it}}) = \beta_{jW}W_{i} + \beta_{jSI}SIL_{i} + \beta_{jc}H_{ome Wealth_{ist-4}} * g(t_{-4}, t_{-1}) + \beta_{jx}X_{it},
\]

where the dependent variable is the log-odds ratio of entry into either incorporated or unincorporated self-employment relative to not switching into self-employment. \( X \) includes the
same controls defined above. The sample includes individuals who were not self-employed in either $t-2$ or $t-4$.

Before turning to the multinomial logit regressions, we begin by assessing—and validating—the “first-stage.” That is, we evaluate whether $\text{Home Wealth}_{it-4} \ast g_{(t-4,t-1)}$ predicts $\text{Home Wealth}_{it}$ after controlling for $\text{Home Wealth}_{it-4}$, $g_{(t-4,t-1)}$, Wages(25-29), and $X$. We conduct these analyses using OLS in columns (1) and (2) of Table IV, where the column (2) regression includes individual fixed effects. As shown, $\text{Home Wealth}_{it-4} \ast g_{(t-4,t-1)}$ enters positively and significantly at the one percent level, when controlling for lagged values of the individual’s home wealth, the recent growth rate of home prices in the state, early career wages, and the array of control variables and fixed effects listed above.

We next examine entry into the aggregate group of self-employed. For these analyses, we use a logit estimator since the dependent variable is a simple one-zero indicator variable. As shown in column (3), $\text{Home Wealth}_{it-4} \ast g_{(t-4,t-1)}$ does not help account for entry into aggregate self-employment. This is consistent with findings that, on average, liquidity constraints do not account for entry into self-employment.

When distinguishing between the incorporated and unincorporated, we discover that collateral impacts entry into entrepreneurship but not into unincorporated self-employment. As shown in columns (4) and (5), $\text{Home Wealth}_{it-4} \ast g_{(t-4,t-1)}$ enters positively and significantly when examining entry into incorporated self-employment but not when considering the odds of switching into unincorporated self-employment. The economic magnitudes are material. For example, consider two similar individuals, where each has $100,000 of home wealth in $t-4$. Let one live in a state where housing prices rise by 25% from $t-4$ to $t-1$ while the other resides in a state where housing prices stagnate. The coefficient estimates indicate that the odds that the individual receiving the positive housing price shocks switches from salaried employment to incorporated self-employment in year $t$ are 4% higher than the otherwise similar individual who did not receive this housing price boost ($1.04=\exp(0.25*0.1566)$).
We were concerned that (a) state housing price growth might be correlated with changes in the state’s overall economic conditions, (b) home wealth is correlated with other individual traits that independently shape entry into entrepreneurship, and (c) these other individual traits are sensitive to overall economic conditions. Under these conditions, $Home\ Wealth_{it-4} \ast g_{(t-4,t-1)}$ might proxy for the interactive impact of (non-home wealth) individual traits and changes in overall economic conditions on entry into entrepreneurship, so that the results in column (6) cannot be interpreted as the impact of collateral on the odds of switching into entrepreneurship. To address this concern, we controlled for shocks to overall economic conditions by including changes in the state unemployment rate (individual $i$’s state) between $t-1$ and $t$ ($\Delta Unemployment$), and its interaction with housing wealth in period $t-4$ ($Home\ Wealth_{it-4} \ast \Delta Unemployment$). The results reported in Table IV hold.

We also conducted a falsification test, similar to the one presented in Table III, to address the concern that $Home\ Wealth(t-4) \ast g(t-4, t-1)$ is capturing something else about an individual besides a shock to home wealth. If $Home\ Wealth(t-4) \ast g(t-4, t-1)$ captures shocks to a person’s collateral between year $t-4$ and $t-1$, then we expect that (a) $Home\ Wealth(t-4) \ast g(t-4, t-1)$ will positively influence selection into incorporated self-employment in year $t$ and (b) $Home\ Wealth(t-4) \ast g(t+1, t+4)$ will not explain entry into entrepreneurship. That is, we would not expect that a shock to future household wealth would influence entry into entrepreneurship unless these future shocks are capturing something else about the evolving characteristics of the individual. In Appendix Table III, we show that $Home\ Wealth(t-4) \ast g(t+1, t+4)$ does not explain entry into entrepreneurship in period $t$. While shocks to wealth before period $t$ explain entry into entrepreneurship, shocks to wealth after period $t$ do not.
V.D. Entry into entrepreneurship: Effective entrepreneurial human capital

Although we have focused on how entrepreneurial ability shapes selection into different employment types, we now expand the notion of entrepreneurial ability from one skill—entrepreneurial ability—to include a second skill: the “capacity,” or willingness, to use entrepreneurial skills to undertake entrepreneurial ventures. The theoretical model with risk aversion developed in Section III.D motivates this extension. The model with risk aversion predicts positive selection into entrepreneurship on the interaction between entrepreneurial ability ($\theta_l$) and noncognitive attitudes toward risk ($\gamma_l$), where these noncognitive attitudes shape the capacity/willingness to use entrepreneurial abilities to start and run a business. Thus, this interaction term represents effective entrepreneurial human capital as a mixture of narrowly defined entrepreneurial ability and the noncognitive traits that that give individuals the capacity to exercise those skills.

Our extended model with risk aversion highlights the conceptual and empirical advantages of using a measure of effective entrepreneurial human capital. In a risk neutral economy, the model suggests that if early career wages are a good proxy for people’s ability to establish a risk-free business, we may find positive selection into entrepreneurship on wages: The best-paid employees turn out to be the most successful business owners. Yet, when risk matters, this prediction does not necessarily hold. Even when early career wages are a good proxy for pure entrepreneurial ability, high wage employees will not necessarily make the most successful business owners. Rather, the positive selection into entrepreneurship on wages should hold only among people with the non-cognitive capacity to “deal with risk” and exercise their entrepreneurial skills in entrepreneurial ventures.

Our model, therefore, predicts that we should examine the interaction between Wages (25-29) and Illicit and that it should be this interaction term—and not Wages (25-29) or Illicit independently—that explains selection into entrepreneurship. The model also predicts that Wages (25-29)*Illicit will not account for selection into other forms of self-employment.
As shown in Table V, the results confirm these predictions. The analyses in Table V are the same as those in Table IV except that we also include $Wages (25-29) \times Illicit$. $Wages (25-29)$ and $Illicit$ were independently included in Table IV and are also included in Table V. There are two key findings. First, we confirm the results from Table IV: (a) shocks to collateral $(Home Wealth_{t-4} \times g(t-4,t-1))$ are positively associated with selection into incorporated self-employment but not into unincorporated self-employment. Second, and consistent with the extended theoretical model, we find that $Wages (25-29)$ is positively associated with selection into incorporated self-employment only among individuals with above the median $Illicit$ scores. These findings suggest that entrepreneurial ability is most strongly associated with selection into entrepreneurship among people with the noncognitive capacity to use those skills in entrepreneurial ventures.

VI. EMPIRICAL RESULTS: CYCLICALITY OF ENTREPRENEURSHIP

The model yields distinct predictions about the cyclicality of entrepreneurship and other self-employment. For example, consider a recession as involving (1) a drop in the demand for salaried workers (labor demand effect) and (2) a tightening of credit conditions (liquidity constraint effect). As discussed above, the model indicates that as labor market opportunities worsen this will have a countercyclical effect on both types of self-employment: A deterioration of labor market opportunities induces more people to sort into both types of self-employment; and a boom in labor market opportunities attracts people out of both forms of self-employment and into salaried jobs. The liquidity constraint effect is different. A tightening of liquidity constraints in a recession discourages entry into entrepreneurship but has no effect on other self-employment that requires no capital. Thus, the model predicts (1) that other self-employment is countercyclical and (2) entrepreneurship is procyclical when the liquidity constraint effect dominates the labor demand market effect. Under this condition, aggregating entrepreneurs and other self-employed individuals will hide the distinctive cyclical patterns of entrepreneurs.
VI.A. The cyclicality of entrepreneurship

To assess empirically the model’s predictions concerning the cyclicality of entrepreneurship and other self-employment—and to evaluate whether self-employment in general provides a misleading perspective on the cyclicality of entrepreneurship, we document the basic cyclical patterns of salaried workers, the aggregate group of self-employed, the incorporated self-employed, and the unincorporated self-employed. To document these patterns, we use state unemployment rates to measure local economic conditions. The Bureau of Labor Statistics produces data on state unemployment for each month. We compute Unemployment as the average unemployment rate in an individual’s state over the twelve months prior to the individual’s interview with the NLSY79.

We estimate the following set linear probability models:

\[ E_{jit} = \beta_j + \beta_{ju} Unemployment_{st} + \beta_{jx} X_{it} + \epsilon_{jist}. \] (14)

\( E_{jit} \) is a binary indicator that equals one if person \( i \) from state \( s \) is observed in employment type \( J \) in time \( t \) and zero otherwise.\(^{10}\) \( Unemployment_{st} \) is the unemployment rate of state \( s \) in year \( t \). \( X_{it} \) is the same set of controls discussed above. We provide the results without (Panel A) and with (Panel B) individual fixed effects.

Table VI reports the coefficient estimates on state unemployment and also gives the mean of the dependent variables. As shown in the column reporting the means of the dependent variables, the proportion of salaried workers, unincorporated self-employed, and incorporated self-employed in our sample are 80.7%, 6.8% and 1.7% respectively.

There are three key findings from Table VI: (1) entrepreneurship is procyclical, (2) unincorporated self-employment is countercyclical, and (3) aggregate self-employment is countercyclical when including individual fixed effects. As shown in Panel B, the state unemployment rate enters negatively and significantly when the dependent variable is

---

\( ^{10} \) The patterns also hold when examining the extensive margin, i.e., when examining the number of hours that individual \( i \) from state \( s \) in year \( t \) works in each employment type \( J \) as the dependent variable.
incorporated self-employment, but enters positively and significantly when the dependent variable is either unincorporated self-employment or the aggregate group of self-employed. Thus, we both confirm the common finding that aggregate self-employment is countercyclical and document that entrepreneurship is procyclical.

The estimated magnitudes are economically large. Consider, the analyses controlling for individual fixed effects. The coefficient estimates indicate that a one-percentage point increase in the state unemployment rate (i.e., an increase of 0.01) is associated with a 1.10% drop in salaried employment relative to the average number of salaried workers \( \frac{1.10\% = [100\times 0.01 \times 0.89]}{0.807} \). The “elasticity” is much larger for incorporated and unincorporated self-employment. Relative to the average number of incorporated and unincorporated self-employed respectively, a one-percentage point increase in the state unemployment rate is associated with 4.16% increase in unincorporated self-employed and a 5.35% decrease in incorporated self-employment.\(^{11}\)

**VI.B. The cyclicality of entrepreneurship: First differences**

In this subsection, we further exploit the longitudinal nature of the NLSY79 to account for omitted state-year factors and draw more confident inferences about the relationship between business cycles and selection into entrepreneurship and other employment types.

In light of these concerns, we estimate the following net entry regressions and report the results in Table VII:

\[
\Delta E_{Jist} = \gamma_J + \gamma_J\Delta Unemployment_{st} + \gamma_J\Delta X_{it} + \nu_{Jist}, \tag{15}
\]

where \( \Delta E_{Jist} \) is the change into employment type \( J \) of individual \( i \) between periods \( t-2 \) and \( t \), so that \( \Delta E_{Jist} \) equals +1 if the individual moves into employment type \( J \); -1 if the person leaves type \( J \); and 0 if the individual does not change designation with respect to employment type \( J \).

\( \Delta Unemployment_{st} \) is the change in the state unemployment rate between year \( t \) and \( t-1 \). Thus, \( \gamma_J\Delta U \) is the coefficient estimate on the relationship between a change in the state’s unemployment

\(^{11}\) For the unincorporated, there is an increase of \( 4.16\% = (100\times 0.01\times 0.283)/0.068 \); and for the incorporated, there is a decrease of \( 5.35\% = (100\times 0.01\times 0.091)/0.017 \).
rate and switches into and out of each employment type. As above, the regressions control for schooling, potential work experience, gender-year, race-year, and state fixed effects. We also provide the results without (Panel A) and with (Panel B) individual fixed effects.

Including state and individual fixed effects in these first differences regressions conditions out many potentially confounding factors. By controlling for state fixed effects, we control for state-specific trends in the probability that individuals select into or out of particular employment types. By controlling for individual fixed effects in these first differences regressions, we not only difference out individual effects, we control for individual trends.

As shown in Table VII, the results from the first differences analyses are consistent with the theoretical model’s predictions and confirm the earlier findings on the cyclicality of entrepreneurship. Entrepreneurship is procyclical and unincorporated self-employment is countercyclical. The results highlight the importance of distinguishing between entrepreneurs and other types of self-employed individuals.

VII. CONCLUSIONS

In this paper, we addressed several gaps that have emerged between theoretical and empirical analyses of entrepreneurship. We begin by offering a new three-sector Roy model of selection into entrepreneurship, other self-employment, and salaried work on human capital and liquidity constraints. The model predicts that (1) entrepreneurs are positively selected on entrepreneurial talent, but the other self-employed are negatively selected on those same skills and traits, (2) entrepreneurs are positively selected on salaried wages—when there is a sufficiently strong connection between entrepreneurial ability and productivity as a salaried worker, but the other self-employed are negatively selected on salaried wages, (3) entrepreneurs are positively selected on collateral, but entry into other forms of self-employment is unrelated to liquidity constraints, and (4) entrepreneurship is procyclical—when business cycles exert a sufficiently strong effect on liquidity constraints, but other forms of self-employment are unambiguously countercyclical. Thus, the model suggests that existing puzzles and unresolved
debates concerning human capital, earnings, liquidity constraints, and the cyclicality of business starts might reflect the failure to distinguish between entrepreneurs and the other self-employed.

Consistent with the theoretical model, we discover that (1) the incorporated are positively selected on proxies for entrepreneurial talent, but the unincorporated are negatively selected on entrepreneurial talent, (2) the incorporated are positively selected on salaried wages, but the unincorporated are negatively selected on wages, (3) collateral exerts a large, positive impact on entry into incorporated self-employment, but collateral does not influence entry into unincorporated self-employment, and (4) entrepreneurship is procyclical, suggesting that business cycles have large liquidity effects, but other forms of self-employment are countercyclical. The results highlight the conceptual and empirical shortcoming of using the aggregate group of self-employed to assess selection into entrepreneurship as human capital and liquidity constraints shape entry into entrepreneurship very differently from entry into unincorporated self-employment.
REFERENCES


<table>
<thead>
<tr>
<th>Human capital</th>
<th>Employed</th>
<th>Total</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFQT</td>
<td>49.2</td>
<td>51.4</td>
<td>48.3</td>
<td>59.8</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>0.05</td>
<td>0.09</td>
<td>-0.02</td>
<td>0.35</td>
</tr>
<tr>
<td>Locus of control</td>
<td>-0.05</td>
<td>-0.21</td>
<td>-0.13</td>
<td>-0.43</td>
</tr>
<tr>
<td>Illicit</td>
<td>-0.02</td>
<td>0.17</td>
<td>0.21</td>
<td>0.07</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>0.18</td>
<td>0.20</td>
<td>0.16</td>
<td>0.28</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>13.8</td>
<td>13.7</td>
<td>13.5</td>
<td>14.4</td>
</tr>
<tr>
<td>College graduate</td>
<td>29%</td>
<td>28%</td>
<td>24%</td>
<td>39%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demographics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>49%</td>
<td>33%</td>
<td>37%</td>
<td>24%</td>
</tr>
<tr>
<td>Black</td>
<td>14%</td>
<td>11%</td>
<td>14%</td>
<td>5%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7%</td>
<td>5%</td>
<td>6%</td>
<td>4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wages:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>2.35</td>
<td>2.39</td>
<td>2.32</td>
<td>2.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wealth</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wealth</td>
<td>$49,939</td>
<td>$94,018</td>
<td>$69,017</td>
<td>$159,763</td>
</tr>
<tr>
<td>Home Wealth</td>
<td>$13,722</td>
<td>$22,982</td>
<td>$19,537</td>
<td>$32,007</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Starting Capital</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting Capital (Mean)</td>
<td>$50,508</td>
<td>$35,715</td>
<td>$90,555</td>
<td></td>
</tr>
<tr>
<td>Starting Capital (Median)</td>
<td>$3,463</td>
<td>$2,033</td>
<td>$19,633</td>
<td></td>
</tr>
<tr>
<td>None needed</td>
<td>17%</td>
<td>21%</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Employees</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Employees (Mean)</td>
<td>2.7</td>
<td>0.7</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>Employees (Median)</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td></td>
</tr>
</tbody>
</table>

Note: The table provides summary statistics on individuals and their businesses while differentiating by whether the person is not a business owners (Employed), a business owners (Total), and if the person is a business owners, the legal form of the business (Unincorporated or Incorporated). The data are from the 2010 and 2012 business ownership part of the NLSY79 survey. Individuals are classified as incorporated or unincorporated only if the legal form of the business from the business ownership part of the NLSY79 survey is confirmed by the individual employment type part of the survey. For the Sources of starting capital, the respondents indicate with each category was an actual component of the capital used to start the business. We examine full-time, full-year individuals. Appendix Table 1 provides variable definitions.
<table>
<thead>
<tr>
<th></th>
<th>Logit (1)</th>
<th>Multinomial Logit (2-3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>-0.1851***</td>
<td>-0.3713*** 0.3139**</td>
</tr>
<tr>
<td></td>
<td>(0.0587)</td>
<td>(0.0653) (0.1491)</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>-0.0861</td>
<td>-0.2683** 0.5198**</td>
</tr>
<tr>
<td></td>
<td>(0.1198)</td>
<td>(0.1324) (0.2628)</td>
</tr>
<tr>
<td>Home Wealth (t-2)</td>
<td>0.0654***</td>
<td>0.0344 0.1607***</td>
</tr>
<tr>
<td></td>
<td>(0.0148)</td>
<td>(0.0246) (0.0213)</td>
</tr>
<tr>
<td>Observations</td>
<td>93,755</td>
<td>93,755 93,755</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.0258</td>
<td>0.0912 0.0912</td>
</tr>
</tbody>
</table>

Notes: This table reports logit (columns 1) and multinomial logit (columns 2-3) analyses of selection into different employment types in year \( t \) on early career salaried wages (Wages (25-29)), Smart & Illicit (which is a zero-one indicator that equals one if the individual had above the median values of AFQT and Illicit in the initial years of the sample), and the net value of the individual's home in year \( t-2 \) (Home Wealth (t-2)). In columns (1), the dependent variable is an indicator variable of whether the individual is self-employed (either unincorporated or incorporated) in year \( t \). Columns (2-3) report the results of multinomial logit regressions, where we do not report the results on unpaid family and other business ownership categories. All regressions include Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Rosenberg self-esteem, Rotter Locus of Control), as well as gender-year, race-year, and state fixed effects. Since the data on home wealth begins in sample year 1985 and we restrict the sample to individuals with data on home wealth in \( t-4 \), the sample starts in 1989. The sample also excludes individuals who were self-employed in either \( t-2 \) or \( t-4 \). Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors, clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
## TABLE III
SELECTION ON WAGES AND HOME WEALTH: OLS

<table>
<thead>
<tr>
<th></th>
<th>Self Employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>-0.0058***</td>
<td>-0.0036</td>
<td>-0.0039</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0037)</td>
<td>(0.0037)</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>-0.0027</td>
<td>-0.0074**</td>
<td>0.0046**</td>
</tr>
<tr>
<td></td>
<td>(0.0039)</td>
<td>(0.0034)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>Home Wealth (t-2)</td>
<td>0.0025***</td>
<td>0.0018**</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0008)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Home Wealth (t+2)</td>
<td></td>
<td>0.0005</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0008)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>Individual FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>93755</td>
<td>93755</td>
<td>88448</td>
</tr>
<tr>
<td>R-square</td>
<td>0.0074</td>
<td>0.2545</td>
<td>0.2672</td>
</tr>
</tbody>
</table>

Notes: This table reports OLS analyses of selection into different employment types in year \( t \) on early career salaried wages (Wages (25-29)), Smart & Illicit (which is a zero-one indicator that equals one if the individual had above the median values of AFQT and Illicit in the initial years of the sample), and the net value of the individual's home in year \( t-2 \) (Home Wealth (t-2)). As a falsification test, columns 3, 6, and 9, examine the net value of the individual's home in year \( t+2 \) (Home Wealth (t+2)). The dependent variable is a one-zero indicator variable of whether the individual is self-employed (columns 1-3), unincorporated self-employed (columns 4-6), or incorporated self-employed (columns 7-9) in year \( t \). All regressions include Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Rosenberg self-esteem, Rotter Locus of Control), as well as gender-year, race-year, and state fixed effects. As indicated, all regressions, except those reported in columns 1, 4, and 7 include individual fixed effects. Since the data on home wealth begins in sample year 1985 and we require values of home wealth in \( t-4 \), the sample starts in sample year 1989. The sample also excludes individuals who were self-employed in either \( t-2 \) or \( t-4 \). The sample if smaller in columns 3, 6, and 9 because the analyses require nonmissing values on Home Wealth in \( t+2 \). Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors, clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
### TABLE IV
SELECTION ON WAGES AND SHOCKS TO HOME WEALTH

<table>
<thead>
<tr>
<th></th>
<th>Home Wealth (OLS)</th>
<th>Home Wealth (OLS-FE)</th>
<th>Self-Employed (Logit)</th>
<th>Unincorporated (Multinominal Logit)</th>
<th>Incorporated (Multinominal Logit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>0.1064***</td>
<td>0.0264</td>
<td>-0.1825***</td>
<td>-0.3637***</td>
<td>0.3143**</td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td>(0.0184)</td>
<td>(0.0585)</td>
<td>(0.0648)</td>
<td>(0.1496)</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>-0.0371</td>
<td>-0.0853</td>
<td>-0.2718**</td>
<td>0.5244**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0310)</td>
<td>(0.1197)</td>
<td>(0.1322)</td>
<td>(0.2629)</td>
<td></td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t-4, t-1)</td>
<td>0.6265***</td>
<td>0.4883***</td>
<td>-0.0001</td>
<td>-0.0711</td>
<td>0.1566**</td>
</tr>
<tr>
<td></td>
<td>(0.0748)</td>
<td>(0.0735)</td>
<td>(0.0482)</td>
<td>(0.0790)</td>
<td>(0.0769)</td>
</tr>
<tr>
<td>g(t-4, t-1))</td>
<td>0.3742***</td>
<td>0.3718***</td>
<td>0.2357</td>
<td>0.4295*</td>
<td>-0.8995*</td>
</tr>
<tr>
<td></td>
<td>(0.0585)</td>
<td>(0.0595)</td>
<td>(0.2057)</td>
<td>(0.2231)</td>
<td>(0.5303)</td>
</tr>
<tr>
<td>Home Wealth (t-4)</td>
<td>0.7575***</td>
<td>0.4993***</td>
<td>0.0650***</td>
<td>0.0177</td>
<td>0.1679***</td>
</tr>
<tr>
<td></td>
<td>(0.0246)</td>
<td>(0.0322)</td>
<td>(0.0163)</td>
<td>(0.0280)</td>
<td>(0.0228)</td>
</tr>
<tr>
<td>Observations</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.5444</td>
<td>0.6475</td>
<td>0.0229</td>
<td>0.0867</td>
<td>0.0867</td>
</tr>
</tbody>
</table>

This table reports analyses of selection into different employment types in year t on Wages (25-29), Smart & Illicit, and a Bartik instrument for changes in home wealth (Home Wealth(t-4)*g(t-4, t-1)), where Home Wealth(t-4) is the individual's net home wealth in year t-4, and g(t-4, t-1) is the growth rate in state housing prices between year t-4 and year t-1 for the state in which the individual lives. Columns (1-2) report OLS regressions in which the dependent variable is the individual’s net home wealth in year t, where column (2) includes individual fixed effects. In column (3), the dependent variable is a one-zero indicator variable of whether the individual is self-employed in year t. In columns (4-5), the dependent variable is a one-zero indicator of employment type, where the reported categories are unincorporated and incorporated respectively, and the unreported categories are salaried, unpaid family, and other business ownership. All regressions include Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Illicit, Rosenberg self-esteem, Rotter Locus of Control), as well as gender, race, year, and state fixed effects. Since the data on home wealth begins in sample year 1985 and we require values of home wealth in t-4, the sample starts in 1989. We exclude individuals who were self-employed in either t-2 or t-4. Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors, clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
### TABLE V

**SELECTION ON WAGES AND SHOCKS TO HOME WEALTH: DIFFERENTIATING BY ILLICIT**

<table>
<thead>
<tr>
<th></th>
<th>Unincorporated (Multinomial Logit)</th>
<th>Unincorporated (Multinomial Logit)</th>
<th>Incorporated (Multinomial Logit)</th>
<th>Incorporated (Multinomial Logit)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>-0.3637***</td>
<td>-0.3653***</td>
<td>0.3143**</td>
<td>0.0703</td>
</tr>
<tr>
<td></td>
<td>(0.0648)</td>
<td>(0.0869)</td>
<td>(0.1496)</td>
<td>(0.1779)</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>-0.2718**</td>
<td>-0.2729*</td>
<td>0.5244**</td>
<td>0.3998</td>
</tr>
<tr>
<td></td>
<td>(0.1322)</td>
<td>(0.1393)</td>
<td>(0.2629)</td>
<td>(0.2739)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t-4, t-1)</td>
<td>-0.0711</td>
<td>-0.0711</td>
<td>0.1566**</td>
<td>0.1541**</td>
</tr>
<tr>
<td></td>
<td>(0.0790)</td>
<td>(0.0790)</td>
<td>(0.0769)</td>
<td>(0.0781)</td>
</tr>
<tr>
<td>g(t-4, t-1))</td>
<td>0.4295*</td>
<td>0.4297*</td>
<td>-0.8995*</td>
<td>-0.8992*</td>
</tr>
<tr>
<td></td>
<td>(0.2231)</td>
<td>(0.2231)</td>
<td>(0.5303)</td>
<td>(0.5295)</td>
</tr>
<tr>
<td>Home Wealth (t-4)</td>
<td>0.0177</td>
<td>0.0177</td>
<td>0.1679***</td>
<td>0.1684***</td>
</tr>
<tr>
<td></td>
<td>(0.0280)</td>
<td>(0.0280)</td>
<td>(0.0228)</td>
<td>(0.0228)</td>
</tr>
<tr>
<td>Wages (25-29)*Illicit</td>
<td>0.0031</td>
<td></td>
<td>0.5479*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1181)</td>
<td></td>
<td>(0.2841)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.0914</td>
<td>0.0915</td>
<td>0.0914</td>
<td>0.0915</td>
</tr>
</tbody>
</table>

This table reports analyses of selection into different employment types in year t on Wages (25-29), Smart & Illicit, the interaction between early career salaried waged and Illicit (Wages (25-29)*Illicit), and a Bartik instrument for changes in home wealth (Home Wealth(t-4)*g(t-4, t-1)), where Home Wealth(t-4) is the individual's net home wealth in year t-4, and g(t-4, t-1) is the growth rate in state housing prices between year t-4 and year t-1 for the state in which the individual lives. The dependent variable is a one-zero indicator of employment type, where the reported categories are unincorporated and incorporated respectively, and the unreported categories are salaried, unpaid family, and other business ownership. All regressions include Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Illicit, Rosenberg self-esteem, Rotter Locus of Control), as well as gender, race, year, and state fixed effects. Since the data on home wealth begins in sample year 1985 and we require values of home wealth in t-4, the sample starts in 1989. We exclude individuals who were self-employed in either t-2 or t-4. Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors, clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
<table>
<thead>
<tr>
<th></th>
<th>Worker</th>
<th>Salaried</th>
<th>Self-employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State Unemployment</strong></td>
<td>-0.782***</td>
<td>-0.962***</td>
<td>0.180</td>
<td>0.268***</td>
<td>-0.088*</td>
</tr>
<tr>
<td>(0.137)</td>
<td>(0.167)</td>
<td>(0.110)</td>
<td>(0.100)</td>
<td>(0.052)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.892</td>
<td>0.807</td>
<td>0.085</td>
<td>0.068</td>
<td>0.017</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>161518</td>
<td>161518</td>
<td>161518</td>
<td>161518</td>
<td>161518</td>
</tr>
<tr>
<td><strong>R-square</strong></td>
<td>0.068</td>
<td>0.041</td>
<td>0.020</td>
<td>0.013</td>
<td>0.015</td>
</tr>
</tbody>
</table>

**Panel B: Employment Type vs. State Unemployment Including Individual Fixed Effects**

<table>
<thead>
<tr>
<th></th>
<th>Worker</th>
<th>Salaried</th>
<th>Self-employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State Unemployment</strong></td>
<td>-0.698***</td>
<td>-0.890***</td>
<td>0.193**</td>
<td>0.283***</td>
<td>-0.091**</td>
</tr>
<tr>
<td>(0.087)</td>
<td>(0.109)</td>
<td>(0.077)</td>
<td>(0.070)</td>
<td>(0.039)</td>
<td></td>
</tr>
<tr>
<td><strong>R-square</strong></td>
<td>0.422</td>
<td>0.438</td>
<td>0.418</td>
<td>0.387</td>
<td>0.367</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>161518</td>
<td>161518</td>
<td>161518</td>
<td>161518</td>
<td>161518</td>
</tr>
</tbody>
</table>

**Notes:** This table reports OLS regression results of employment types, Worker, Salaried, Self-employed, Unincorporated, and Incorporated, on state unemployment. Panels A - B each reports the results of six OLS regressions, one for each employment type, where the dependent variable equals 1 if the person has the designated employment type in period $t$ and 0 otherwise. In Panel B, the regressions control for individual fixed effects. Though not shown, all regressions control for schooling (measured in six categories), potential work experience (quartic), gender, race, state, and year-gender fixed effects, and lagged values of the dependent variable. The table also provides the means of the dependent variables. The sample includes who are least 25 years old. Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors clustered at the state-year level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
# TABLE VII
EMPLOYMENT TYPES AND STATE UNEMPLOYMENT: FIRST DIFFERENCES

<table>
<thead>
<tr>
<th></th>
<th>Worker</th>
<th>Salaried</th>
<th>Self-employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: ∆Employment Type vs. ∆State Unemployment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆State Unemployment</td>
<td>-0.628***</td>
<td>-0.919***</td>
<td>0.291**</td>
<td>0.441***</td>
<td>-0.150**</td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td>(0.193)</td>
<td>(0.143)</td>
<td>(0.132)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Observations</td>
<td>160108</td>
<td>160108</td>
<td>160108</td>
<td>160108</td>
<td>160108</td>
</tr>
<tr>
<td>R-square</td>
<td>0.409</td>
<td>0.311</td>
<td>0.195</td>
<td>0.209</td>
<td>0.156</td>
</tr>
<tr>
<td>Panel B: ∆Employment Types vs. ∆State Unemployment: Individual Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆State Unemployment</td>
<td>-0.615***</td>
<td>-0.912***</td>
<td>0.297**</td>
<td>0.456***</td>
<td>-0.159**</td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.180)</td>
<td>(0.132)</td>
<td>(0.122)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>R-square</td>
<td>0.514</td>
<td>0.438</td>
<td>0.340</td>
<td>0.347</td>
<td>0.283</td>
</tr>
</tbody>
</table>

**Notes:** This table reports OLS regression results of the change in employment type (Worker, Salaried, Self-employed, Unincorporated, and Incorporated) on the change in the state unemployment rate and a set of control variables. The dependent variable equals +1 if the individual moves into the indicated employment type between $t-2$ and $t$; -1 if the person leaves the employment type; and 0 if the individual does not change designation with respect to the indicated employment type. Each panel reports the results of six OLS regressions, one for each employment type. Though not shown, all regressions control for schooling (measured in six categories), potential work experience (quartic), state, year-race, year-gender fixed effects, and lagged values of the dependent variable. The regressions in Panel B also control for individual fixed effects. Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
1. Human capital

**AFQT**

Armed Forces Qualifications Test score measures the aptitude and trainability of the respondent. Collected during the 1980 NLSY79 survey, the AFQT score is based on arithmetic reasoning, world knowledge, paragraph comprehension, and numerical operations. It is frequently employed as a general indicator of cognitive skills. This AFQT score is measured as a percentile of the NLSY79 survey, with a median value of 50.

**Illicit**

Illicit measures the aggressive, risk-taking, disruptive, "break-the-rules," behavior of individuals based on the 1980 NLSY79 survey. Taken from Levine and Rubinstein (2017), this index is based on 20 questions, where 17 concern delinquency, e.g., damaging property, fighting at school, shoplifting, robbery, using force to obtain things, assault, threatening to assault, drug use, dealing drugs, gambling, and so forth, and three are about interactions with the police, e.g., stopped by the policy, charged with an illegal activity, or convicted for activities other than minor traffic violations. For each question, a value of one is assigned if the person responds in 1980 that they engaged in that activity and zero otherwise. The average of the answers is then computed for each individual. Finally, we construct a standardized version by subtracting the sample mean and dividing by the standard deviation to create a mean zero, standard deviation of one indicator of illicit activity.

**Smart & Illicit**

Smart & Illicit equals one if the individual's AFQT score is greater than or equal to 50 and Illicit is greater than or equal to zero and Smart & Illicit equals zero otherwise.

**Rosenberg self-esteem (standardized)**

Rosenberg Self-Esteem score is based on a ten-part questionnaire given to all NLSY79 participants in 1980. It measures the degree of approval or disapproval of one’s self. The values range from six to 30, where higher values signify greater self-approval. Rosenberg Self-Esteem (standardized) standardizes the score, so that it has a mean of zero and a standard deviation of one.

**Rotter locus of control (standardized)**

Rotter Locus of Control measures the degree to which respondents believe they have internal control of their lives through self-determination relative to the degree that external factors, such as chance, fate, and luck, shape their lives. It was collected as part of a psychometric test in the 1979 NLSY79 survey. The Rotter Locus of Control ranges from 4 to 16, where higher values signify less internal control and more external control. This is standardized, so that it has a mean of zero and a standard deviation of one.

**Years of schooling**

The respondent’s maximum number of years of schooling, so it does not vary over time for a respondent.

---

### APPENDIX TABLE I:

**VARIABLE DEFINITIONS AND SOURCES**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Human capital</td>
<td></td>
</tr>
<tr>
<td><strong>AFQT</strong></td>
<td>Armed Forces Qualifications Test score measures the aptitude and trainability of the respondent. Collected during the 1980 NLSY79 survey, the AFQT score is based on arithmetic reasoning, world knowledge, paragraph comprehension, and numerical operations. It is frequently employed as a general indicator of cognitive skills. This AFQT score is measured as a percentile of the NLSY79 survey, with a median value of 50.</td>
</tr>
<tr>
<td><strong>Illicit</strong></td>
<td>Illicit measures the aggressive, risk-taking, disruptive, &quot;break-the-rules,&quot; behavior of individuals based on the 1980 NLSY79 survey. Taken from Levine and Rubinstein (2017), this index is based on 20 questions, where 17 concern delinquency, e.g., damaging property, fighting at school, shoplifting, robbery, using force to obtain things, assault, threatening to assault, drug use, dealing drugs, gambling, and so forth, and three are about interactions with the police, e.g., stopped by the policy, charged with an illegal activity, or convicted for activities other than minor traffic violations. For each question, a value of one is assigned if the person responds in 1980 that they engaged in that activity and zero otherwise. The average of the answers is then computed for each individual. Finally, we construct a standardized version by subtracting the sample mean and dividing by the standard deviation to create a mean zero, standard deviation of one indicator of illicit activity.</td>
</tr>
<tr>
<td><strong>Smart &amp; Illicit</strong></td>
<td>Smart &amp; Illicit equals one if the individual's AFQT score is greater than or equal to 50 and Illicit is greater than or equal to zero and Smart &amp; Illicit equals zero otherwise.</td>
</tr>
<tr>
<td><strong>Rosenberg self-esteem (standardized)</strong></td>
<td>Rosenberg Self-Esteem score is based on a ten-part questionnaire given to all NLSY79 participants in 1980. It measures the degree of approval or disapproval of one’s self. The values range from six to 30, where higher values signify greater self-approval. Rosenberg Self-Esteem (standardized) standardizes the score, so that it has a mean of zero and a standard deviation of one.</td>
</tr>
<tr>
<td><strong>Rotter locus of control (standardized)</strong></td>
<td>Rotter Locus of Control measures the degree to which respondents believe they have internal control of their lives through self-determination relative to the degree that external factors, such as chance, fate, and luck, shape their lives. It was collected as part of a psychometric test in the 1979 NLSY79 survey. The Rotter Locus of Control ranges from 4 to 16, where higher values signify less internal control and more external control. This is standardized, so that it has a mean of zero and a standard deviation of one.</td>
</tr>
<tr>
<td><strong>Years of schooling</strong></td>
<td>The respondent’s maximum number of years of schooling, so it does not vary over time for a respondent.</td>
</tr>
</tbody>
</table>
College graduate  Graduated from college or obtained an advanced degree.

Educational Attainment  The six educational attainment categories: (i) high school dropouts: less than 12 years of schooling (ii) GED degree (iii) high school graduates: 12 years of schooling (iv) had some college education: 13-15 years of schooling (v) college education: 16 years of schooling (vi) advanced studies: 17+ years of schooling. These are measured at the end of the respondent’s educational experience, so that they do not vary over time for a respondent.

Potential Experience  Age of the respondent minus the years of schooling minus six, or, if this computation is less than zero, then potential experience set equal to zero.

Female  Equals one if the respondent reports being female and zero otherwise.

Black  Equals one if the respondent reports being Black and zero otherwise.

Hispanic  Equals one if the respondent reports being Hispanic and zero otherwise.

2. Collateral, Wealth, and Earnings

Home Wealth  The market value of the respondent’s home net of any mortgages.

Wealth  Created by summing all asset values and subtracting all debts.

Wages (25-29)  When the respondent if 31 or more years old, Wages (25-20) equals the respondent's average log real wages (2010 prices) as a salaried employee when the respondent is 25-29 years old. When the respondent is 27-30 years old, Wages (25-29) equals the individual’s average log real hourly earnings as a salaried employee at the age of t-2.

Earnings  Wages plus income from business. Deflated by the CPI corresponding to when those earnings were realized. Earnings are in 2010 prices.

3. Employment Types

Unincorporated  If a respondent is self-employed, the NLSY79 further asks whether the business is incorporated or not. If the respondent is self-employed and the business is unincorporated, then Unincorporated Self-employed equals one and it is zero otherwise.

Incorporated  If a respondent is self-employed, the NLSY79 further asks whether the business is incorporated or not. If the respondent is self-employed and the business is incorporated, then Incorporated Self-employed equals one and it is zero otherwise. See Levine and Rubinstein (2017) for additional coding.

Self-employed  From the NLSY79’s unified class of worker (R24455.10), there are four responses for working respondents: (1) Private company, including non-profit, (2) government, (3) self-employed, and (4) those working without pay, including in family businesses. We set Self-employed equal to one if the respondent’s class of worker is “(3)” and zero otherwise.

Salaried  From the NLSY79’s unified class of worker (R24455.10), there are four responses for working respondents: (1) Private company, including non-profit, (2) government, (3) self-employed, and (4) those working without pay, including in family businesses. We set Salaried equal to one if the respondent’s class of worker is either “(1)” or “(2)” and zero otherwise.

Unpaid family business  Equals one if the respondent indicates that they are unpaid and work in a family business and zero otherwise.
4. Legal Form of Business

**Unincorporated (B)**

Equals one if the respondent indicates that the legal form of the business is a sole proprietorship and zero otherwise. This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

**Incorporated (B)**

Equals one if the respondent indicates that the legal form of the business is either (a) a partnership or limited liability partnership, (b) a limited liability corporation, (c) a sub-chapter S corporation, or (d) a general corporation and zero otherwise. This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

**Other Business (B)**

Equals one if the respondent indicates that the legal form of the business is either (a) a nonprofit organization or (b) other and zero otherwise. This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

6. State-Year Characteristics

**g(x,y)**

The growth rate in state housing prices between years x and y. The Federal Housing Finance Agency provides house price indices by state and year.

**ΔUnemployment**

The change in the unemployment rate in the respondent's state over the preceding twelve months. In particular, the Bureau of Labor Statistics produces data on state unemployment for each month. The NLSY79 gives the date when each person was sampled. We compute change in the state's unemployment over the preceding twelve months from the date of the interview.

---

Notes:

(1) All data are from the NLSY79 unless otherwise indicated.

(2) The NLSY79 is a representative survey of 12,686 individuals who were 15-22 years old when they were first surveyed.

(3) We use the sampling weights provided by the NLSY79.

(4) In Table 1, which covers the survey years 2010 and 2012, we classify an individual as incorporated if both Incorporated and Incorporated (B) indicate that the individual is an incorporated business owner for the 2010 (2012) survey. We get very similar results if we instead use only the business survey (Incorporated (B) to classify the legal form of the business. The same holds for unincorporated business owners.
APPENDIX TABLE II
EARNINGS BY EMPLOYMENT TYPE AND EARLY SALARIED WAGES

Panel A: Earnings vs. early salary wages, standard controls

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Hourly Earnings (31+)</td>
<td>0.510***</td>
<td>0.2367*</td>
<td>0.067</td>
<td>0.936***</td>
</tr>
<tr>
<td>(Observations)</td>
<td>41015</td>
<td>3488</td>
<td>2568</td>
<td>920</td>
</tr>
<tr>
<td>(R-square)</td>
<td>0.101</td>
<td>0.0732</td>
<td>0.102</td>
<td>0.104</td>
</tr>
</tbody>
</table>

Notes: This table provides regression results of log hourly earnings in year $t$ on an individual's average log wages as a salaried employee during the ages of 25 through 29 (Wages (25-29)). All regressions include "standard controls:" Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Rosenberg self-esteem, Rotter Locus of Control), as well as race, gender, year, and state fixed effects. In Panel B, the regressions also include state-year fixed effects. As indicated, each regression includes the subsample of individuals who are salaried (columns 1-5), self-employed (columns 2-6), unincorporated self-employed (columns 3-7), or incorporated self-employed (columns 4-8) in year $t$. The sample includes full-time, full-year workers who are 31 years of age or older. Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors, clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
### APPENDIX TABLE III

**SELECTION ON WAGES AND SHOCKS TO HOME WEALTH:**

**FALSIFICATION TEST**

<table>
<thead>
<tr>
<th></th>
<th>Self-Employed (Logit)</th>
<th>Unincorporated (Multinomial Logit)</th>
<th>Incorporated (Multinomial Logit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>-0.1831***</td>
<td>-0.3670***</td>
<td>0.3162**</td>
</tr>
<tr>
<td></td>
<td>(0.0587)</td>
<td>(0.0651)</td>
<td>(0.1497)</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>-0.0865</td>
<td>-0.2705**</td>
<td>0.5254**</td>
</tr>
<tr>
<td></td>
<td>(0.1197)</td>
<td>(0.1323)</td>
<td>(0.2625)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t, t+4)</td>
<td>0.0818</td>
<td>0.1660</td>
<td>0.0719</td>
</tr>
<tr>
<td></td>
<td>(0.0773)</td>
<td>(0.1113)</td>
<td>(0.1248)</td>
</tr>
<tr>
<td>g(t, t+4)</td>
<td>0.0619</td>
<td>-0.1745</td>
<td>1.3717***</td>
</tr>
<tr>
<td></td>
<td>(0.2437)</td>
<td>(0.2751)</td>
<td>(0.5323)</td>
</tr>
<tr>
<td>Home Wealth (t-4)</td>
<td>0.0618***</td>
<td>0.0069</td>
<td>0.1737***</td>
</tr>
<tr>
<td></td>
<td>(0.0175)</td>
<td>(0.0301)</td>
<td>(0.0252)</td>
</tr>
<tr>
<td>Observations</td>
<td>93722</td>
<td>93755</td>
<td>93755</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.0258</td>
<td>0.0914</td>
<td>0.0914</td>
</tr>
</tbody>
</table>

This table reports analyses of selection into different employment types in year $t$ on Wages (25-29), Smart & Illicit, and a Bartik instrument for changes in home wealth (Home Wealth(t-4)*g(t, t+4)), where Home Wealth(t-4) is the individual's net home wealth in year $t-4$, and g(t, t+4) is the growth rate in state housing prices between year $t+1$ and year $t+4$ for the state in which the individual lives. In column (1), the dependent variable is a one-zero indicator variable of whether the individual is self-employed in year $t$. In columns (2-3), the dependent variable is a one-zero indicator of employment type, where the reported categories are unincorporated and incorporated respectively, and the unreported categories are salaried, unpaid family, and other business ownership. All regressions include Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and non-cognitive traits (AFQT, Illicit, Rosenberg self-esteem, Rotter Locus of Control), as well as gender, race, year, and state fixed effects. Since the data on home wealth begins in sample year 1985 and we require values of home wealth in $t-4$, the sample starts in 1989. We exclude individuals who were self-employed in either $t-2$ or $t-4$. Appendix Table 1 provides variable definitions. Heteroskedasticity robust standard errors, clustered at the individual level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
APPENDIX TABLE IV
EMPLOYMENT TYPES AND HOURS OVER THE BUSINESS CYCLE: CPS

<table>
<thead>
<tr>
<th>Worker</th>
<th>Salaried</th>
<th>Self-employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Employment Type vs. State Unemployment: OLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Unemployment</td>
<td>-0.530***</td>
<td>-0.610***</td>
<td>0.080***</td>
<td>0.113***</td>
</tr>
<tr>
<td>(0.042)</td>
<td>(0.047)</td>
<td>(0.026)</td>
<td>(0.023)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.843</td>
<td>0.756</td>
<td>0.087</td>
<td>0.059</td>
</tr>
<tr>
<td>Observations</td>
<td>2199569</td>
<td>2199569</td>
<td>2199569</td>
<td>2199569</td>
</tr>
<tr>
<td>R-square</td>
<td>0.076</td>
<td>0.034</td>
<td>0.030</td>
<td>0.016</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Employment Type vs State Unemployment: Multinomial Logit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Working</td>
</tr>
<tr>
<td>State Unemployment</td>
</tr>
<tr>
<td>(0.354)</td>
</tr>
</tbody>
</table>

Notes: This table reports OLS and multinomial logit regression results of each employment type (Worker, Salaried, Self-employed, Unincorporated, and Incorporated) on state unemployment using the CPS. Panel A reports the results of five OLS regressions, one for each employment type. The dependent variable is the proportion of individuals in the specified employment type. Panel C reports multinomial logit regression results, where the dependent variable is the log-odds of being in the indicated employment type rather than a salaried worker. Though not shown, all regressions control for race, schooling (measured in six categories), potential work experience (quartic), state fixed effects, and year-gender fixed effects. Panel A also report the means of the dependent variables. Heteroskedasticity robust standard errors clustered at the state-year level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
Figure I: Selection into Employment Types by Entrepreneurial Ability

Figure I illustrates the relationship between the log utility in each employment type and the log of entrepreneurial ability ($\ln \theta$). The horizontal line represents the log utility of other self-employment ($\ln V_{ui}$) and equals $\varepsilon_{ui} + \delta_{ui}$. The upward sloping line with squares is the log utility of salaried employment ($\ln V_{si}$), where the slope is $\rho_S$. $\ln V_{si}$ intersects $\ln V_{ui}$ at the first cutoff level of entrepreneurial ability: $\ln \theta^*_S$. The upward sloping line with circles is the log utility of entrepreneurship ($\ln V_{Ei}$), where the slope is $\rho'_E$, and where $\ln V_{Ei}$ intersects $\ln V_{si}$ at the second cutoff level: $\ln \theta^*_E$. 
Figure II illustrates the change in selection into each employment type when credit conditions tighten, i.e., when $r_l$ increases. The horizontal line represents the log utility of other self-employment ($lnV_{ui}$); the upward sloping line with squares is the log utility of salaried employment ($lnV_{si}$), where the slope is $\rho_S$, and the upward sloping line with circles is the log utility of entrepreneurship ($lnV_{ei}$), where the slope is $\rho_E$. $lnV_{si}$ intersects $lnV_{ui}$ at the first cutoff level of entrepreneurial ability: $ln\theta^*_E$. $lnV_{ei}$ intersects $lnV_{si}$ at the second cutoff level: $ln\theta^{**}_E$. When credit conditions tighten this shifts downward the intercept of the line for the log utility of entrepreneurship, constraining entry into entrepreneurship. Changes in $r_l$, however, do not alter the intercepts or slopes of the other lines and therefore liquidity conditions do not affect entry into other self-employment.
Figure III illustrates the change in selection into each employment type when there is a recession that tightens credit conditions and reduces the demand for salaried workers. The tightening of liquidity conditions involves a parallel fall in the log utility of entrepreneurship line ($\ln \theta^*_E$), i.e., the upward sloping line with circles. As shown, this tightening reduces selection into entrepreneurship but has no effect on entry into other self-employment, i.e., the liquidity effect exerts a procyclical influence on entrepreneurship, but not on other self-employment. The reduction in the demand for salaried workers involves a parallel fall in the log utility of salaried employment ($\ln V_{s1}$) line, i.e., the upward sloping line with squares. This labor demand effect exerts a countercyclical influence on both entrepreneurship and other self-employment. The figure depicts the special case when liquidity and labor demand effects exactly counterbalance each other with respect to entrepreneurship.