ENTREPRENEURSHIP, HUMAN CAPITAL, AND LIQUIDITY CONSTRAINTS*

ROSS LEVINE AND YONA RUBINSTEIN

February 2018

Abstract: Although theory highlights the roles of human capital and liquidity constraints in shaping selection into entrepreneurship, empirical research finds that (a) salaried and self-employed individuals have similar human capital and (b) liquidity constraints bind for very few. To address this gap between theory and evidence, we develop a model in which we differentiate between entrepreneurship and other forms of self-employment. The model predicts—and our empirical analyses confirm—that (1) entrepreneurs are positively selected on salaried wages and associated human capital, but other business owners are negatively selected on those factors and (2) entrepreneurs are positively selected on collateral, but not other business owners. Moreover, we exploit exogenous variation in home equity and discover that collateral impacts entry into entrepreneurship, but not other forms of self-employment.

Keywords: Entrepreneurship; Self-employment; Human capital; Occupational choice; Corporate finance

JEL Classifications: L26; J24; G32

* Levine: University of California, Berkeley, and the NBER, rosslevine@berkeley.edu. Rubinstein: London School of Economics, CEP, and the CEPR, y.rubinstein@lse.ac.uk. We thank David De Meza, Paul Gertler, Erik Hurst, Chinhui Juhn, Ed Lazear, and seminar participants at the London School of Economics, University of California-Berkeley, and the University of Houston.
I. INTRODUCTION

Entrepreneurship plays a central role in influential theories of innovation, economic growth, business cycles, and firm productivity. For example, Smith (1776), Schumpeter (1911), and Aghion and Howitt (1992) all emphasize that entrepreneurs facilitate economic growth by bring new goods, services, and production processes to the economy. In Fisher (1933), Keynes (1936), Bernanke and Gertler (1989), Shleifer (1986), and Caballero and Hammour (1994), the response of entrepreneurs is crucial to understanding the amplification and propagation of aggregate shocks. And, in Lucas (1978), Baumol (1990), Murphy et al (1991), and Gennaioli et al (2013), entrepreneurial human capital helps determine the productivity of firms and the growth rates of regional and national economies. Unsurprisingly, therefore, a rich body of research explores the factors determining selection into entrepreneurship.

Many theories emphasize that human capital and liquidity constraints shape entry into entrepreneurship. Schumpeter (1991), Lucas (1978), Kihlstrom and Laffont (1979), Evans and Jovanovic (1979), Baumol (1990), Murphy et al (1991), Gennaioli et al (2013), and others stress that unique human capital abilities—including creativity, analytical skills, risk taking, confidence, education, and managerial and marketing acumen—influence who becomes and succeeds as an entrepreneur. An extensive body of theoretical research also stresses that access to capital is decisive in determining entry into entrepreneurship, e.g., Knight (1921), Bernanke and Gertler (1989), Evans and Jovanovic (1989), and Kyotaki and Moore (1997). If starting a business requires a minimum amount of capital, then individuals need external financing if they are unable or unwilling to self-finance the business and this implies that liquidity constraints materially influence selection into entrepreneurship.

Empirical research, however, highlights two corresponding puzzles about entrepreneurship, human capital, and liquidity constraints. First, although influential theories emphasize the unique human capital attributes of entrepreneurs, the data emphasize the human capital similarities between business owners and salaried workers. As summarized below, they
have similar education, learning aptitude scores, personality characteristics, and earnings. The second puzzle regards liquidity constraints. As we document below, most U.S. businesses were started with less than $3,500 of capital, a fifth started with zero capital, and less than 5% obtained external credit from a financial institution to start the business. Although researchers discover a significant relationship between collateral and entry into self-employment,\(^1\) Hurst and Lusardi (2004) show that liquidity constraints bind for only a very small proportion of the wealthiest business owners. Given the centrality of liquidity constraints in theoretical explanations of entry into entrepreneurship, it is puzzling both that businesses require little starting capital and that liquidity constraints seem to bind for so few.

Motivated by these puzzles, we develop a theoretical model and provide empirical evidence that shed new light on how human capital and liquidity constraints shape selection into entrepreneurship. A key starting point is the growing body of evidence that business ownership 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), and Levine and Rubinstein (2017) indicate that while some of the self-employed undertake highly-productive ventures that create jobs and introduce new goods and services to the market, many are one-person, low-productivity business owners who were unsuccessful salaried workers, perform routine, manuals tasks, and have no ambitions to grow their businesses. Thus, studying the self-employed might yield misleading inferences about entrepreneurs. Indeed, we show that distinguishing between entrepreneurs and other self-employed individuals can resolve both the human capital and liquidity puzzles.

We first develop a three-sector Roy model that distinguishes between entrepreneurs, salaried employees, and the non-entrepreneurial self-employed, which we call self-employed. Our model differs from Evans and Jovanovic’s (1989) (EJ) influential model of entrepreneurship.

in a key respect. While EJ aggregate business owners into one category of self-employment, we distinguish between entrepreneurship—which demands entrepreneurial ability, physical capital, and liquidity—and other forms of self-employment that demand none of these. Our model also differs from EJ in that we relax their assumption that entrepreneurial ability is uncorrelated with productivity as a salaried employee. We allow entrepreneurial ability to positively affect earnings in paid 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 two unique predictions with respect to human capital and liquidity constraints. First, ceteris paribus, entrepreneurs are positively selected on entrepreneurial ability and salaried wages, but the self-employed are negatively selected on these same factors. Thus, aggregating the high-ability, high-wage entrepreneurs with the low-ability, low-wage self-employment may yield a composite group of business owners that has, on average, similar human capital traits to wage earnings, which accounts for the human capital puzzle. The model’s second prediction is that entrepreneurs are negatively selected on liquidity constraints, but the self-employed are not. Thus, combining these two types of business owners may yield an aggregate group in which only a small proportion is liquidity constrained.

In turning to the data, we follow Levine and Rubinstein (2017) (LR) and use the incorporated as a proxy for “entrepreneurs” and the unincorporated as a proxy for the model’s self-employed. 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, LR 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.

Using the National Longitudinal Survey of Youths 1979, we document that the incorporated and unincorporated are notably different with respect to human capital, starting capital, and the sources of 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 business owners. Furthermore, when comparing the salaried wages of people when they were in their 20s, those who become incorporated self-employed earned more as salaried workers early in their careers than those who either remained salaried workers or switched into unincorporated business ownership. There are also startlingly 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. In terms of the sources of starting capital, the incorporated are much more likely to receive financing through the financial system. For example, the incorporated are more than three times more likely to receive a bank loan, four times more likely to receive external financing from a non-bank (non-governmental) entity, and six-times more likely to use a home equity loan to start their businesses.

To assess the model’s predictions, we evaluate the differential selection of individuals into incorporated and unincorporated self-employment on (1) salaried wages and associated human capital and (2) collateral. We first discover that entrepreneurs—as proxied by the incorporated self-employed—are positively selected on salaried wages, while the unincorporated are negatively selected on wages. 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 and liquidity puzzles: When researchers combine entrepreneurs with other business owners, this
aggregate away the human capital differences between entrepreneurs and other employment
types and obfuscates the connection between entrepreneurship and liquidity constraints.

Finally, 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 self-employment. 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.

II. THE HUMAN CAPITAL AND LIQUIDITY PUZZLES

In this section, we document two puzzles emerging from the literature on
entrepreneurship. This first is that several human capital indicators suggest that salaried
employees and self-employed business owners are very similar despite an abundance of
theoretical models emphasizing the distinct features of entrepreneurs. Second, although many
theoretical models assume that liquidity constraints severely constrain entry into
entrepreneurship, most businesses start with very little capital. To illustrate these puzzles—and
foreshadow our strategy for resolving them, we use data from the NLSY79.

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.

2 The Data Appendix provides detailed variable definitions and sources.
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 property, fighting, shoplifting, robbery, assault, drug use and dealing, etc., and whether the individual was stopped by the policy, 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, following Levine and Rubinstein (2017), 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 Rosenberg Self-Esteem index, which is based on ten questions in the 1980 survey, measures the degree of approval or disapproval of one’s self and has been widely used in psychology and economics. Rotter 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 Rotter 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 and earnings. 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 at the age of $t-2$ 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 more detailed questions about businesses, including the amount of capital used to start the business, the source of that capital, and the number of employees. In terms of the sources of starting capital, owners might use family savings, external finance (such as home equity loans, bank loans, loans from businesses, etc.), or other sources (such as credit card debt, or government loans).

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 identify, 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 self-employed in our model. As emphasized in the Introduction, 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 and their businesses 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 business owners. For business owners, the table provides summary statistics on all business owners (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.\(^3\)

Table I shows that the human capital of employees and business owners is remarkably similar. Employees and business owners in 2010 and 2012 have, on average, virtually the same number of years of education (13.8 v. 13.7), and the same proportion of college graduates (29% v. 28%). Since wages reflect human capital, we consider people who are employees and business owners in 2010 and 2012 and compare their salaried earnings when they were 25-29 year old. As shown, employees and business owners in 2010 and 2012 had almost identical salaried wages, on average, when they were 25-29 years old: 2.35 for employees and 2.39 for business owners. 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 business owners are small, though business owners have slightly higher AFQT scores and self-esteem values, and slightly lower Locus of control measures. For example, there is only a 2.2 percentile point difference in average AFQT scores between employees (49.2) and business owners (51.4). Thus, although influential models of entrepreneurship emphasize the unique human capital of entrepreneurs, business owners 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 business owners, those who tend to engage in entrepreneurial activities (Incorporated) and those who do not (Unincorporated), and these two types of business owners have very

---

\(^3\) There are some differences between the responses that individuals give regarding their employment type in the business ownership and employment parts of the NLSY79. In Table 1, we classify an individual as an incorporated or unincorporated business owner 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 of the survey.
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 unincorporated business owners 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 business owners, entrepreneurs have much greater human capital, on average, than salaried workers, while other business owners have much less. Aggregation may account for the human capital puzzle.

II.C. Patterns: Starting capital

Table I also documents the liquidity puzzle: Most businesses start with very little capital. The median value of starting capital is less than $3,500 and most businesses do not have any 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 owners indicate that no capital was needed to start their businesses. Furthermore, among all business owners, only 11% indicate that they received any external financing, e.g., home equity loans, banks loans, or other types of business loans. These observations raise questions about whether liquidity constraints represent a high entry barrier for the median business.

The notable differences between entrepreneurs and other business owners 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. While 6% of the unincorporated self-employed use external financing to start their businesses, 24% of the incorporated received loans to start their business. There are also pronounced differences in wealth. The total wealth of unincorporated business owners is, on average, about $70,000, where $19,500 is home wealth. In contrast, the overall wealth of incorporated business owners is almost $160,000, with $32,000 of that in the form of home wealth. These notable differences in collateral, starting capital, and the sources of 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.

These summary statistics illustrate differences between the starting capital of entrepreneurs and other business owners, but they do not consider the potential correlation across different sources of capital. Thus, we now extend these analyses with multivariate regressions to provide additional information on the relationship between the source of starting capital and both the legal form of the business and the amount of starting capital. Table II has two sets of the regression. In columns (1) - (3), we examine the relationship between the sources of capital and the legal form of the business. When using a linear probability model (column 1) or a probit model (column 2), the dependent variable equals one if the business is incorporated and zero otherwise. When using a multinomial logit (ML) specification (column 3), we report the results on incorporated self-employment, where the unreported categories for the business's legal form are unincorporated and other legal forms. In the second set of regressions, columns (4) – (5), we examine the relationship between the sources and value of starting capital. We provide both ordinarily least squares (column 4) and median estimates (column 5). In all specifications, the explanatory variables are the one-zero indicator variables of each source of capital.

Two messages emerge from Table II. First, entrepreneurs are much more likely to use home equity loans, bank loans, and business loans to start their businesses than other business owners. These results hold when using OLS, probit, and ML regressions. For example, the
Parameter estimates from the linear probability model (column 1) indicate that incorporated businesses are 28% more likely to have received a home equity loan and about 23% more likely to have received a bank loan than other businesses. The second message is that there is a strong link between the source of starting capital and the amount of starting capital. For example, the parameter estimates from the OLS regression (column 4) show that businesses receiving bank loans as part of their starting capital begin, on average, with $180,000 more capital than other businesses and (2) those using home equity loans start business begin, on average, with $142,500 more starting capital than other business owners.

The differences between entrepreneurs and other business owners with respect to starting capital are even more pronounced when analyzing quartiles. Figure I depicts the amount of starting capital at each quartile of the starting capital distribution. This is done separately for the aggregate group of self-employed, the incorporated, and the unincorporated. Thus, the starting capital of the fourth quartile of the incorporated self-employment distribution is almost $200,000, and the starting capital for the fourth quartile of the unincorporated distribution is about $50,000. The figure shows that for three-quarters of the distribution of unincorporated businesses owners, exceedingly little capital was used to start their businesses. This further emphasizes the potential pitfalls of aggregating the incorporated and unincorporated.

Quantile illustrations also help motivate our focus on the complementarities between human and physical capital. In Figure II, we again examine the starting capital quartiles for all business owners, the incorporated, and the unincorporated. The figure gives AFQT scores for each self-employment category at those quartiles. As shown, the AFQT scores of incorporated business owners are above the sample median of 50 at each starting capital quartile. Moreover, the AFQT level at the fourth quartile of the starting capital distribution of incorporated business owners is an extraordinarily high 63. People initiating entrepreneurial activities with large injections of starting capital have exceptionally high AFQT scores, highlighting the strong association among AFQT, liquidity, and entrepreneurship. This does not hold for the
unincorporated. The AFQT scores of unincorporated business owners are below the sample median at each quartile.

II.D. Discussion

Tables I-II and Figures I-II both document two puzzles and suggest a strategy for resolving those puzzles. With respect to the puzzles, salaried employees and self-employed business owners have similar human capital characteristics and most businesses start with little or no capital. With respect to resolving these puzzles, the data suggest that there are material differences between incorporated and unincorporated business owners and their businesses. Thus, we now develop a three-sector Roy model to explore the selection of individuals on human capital and liquidity into entrepreneurship and other forms of self-employment. Below, we empirically evaluate the predictions emerging from the model.

III. MODEL

III.A. Framework

Each individual chooses one of three employment types: Self-employment \((U)\), salaried employment \((S)\), or entrepreneurship \((E)\). Individual \(i\) then receives income \(I_{ji}\) from working in employment type \(J\), where \(J\) is \(U\), \(S\), or \(E\). Individuals sort into employment types to maximize utility, where the utility of individual \(i\) in employment type \(J\) \((V_{ji})\) is a function of income and non-pecuniary benefits \((\delta_{ji})\):\(^4\)

\[
V_{ji} = I_{ji} * e^{\delta_{ji}},
\]

where 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 EJ, and then show below that the core predictions hold when allowing for risk aversion.

\(^4\) For example, the non-pecuniary benefits of self-employment could reflect preferences to be one’s “own boss,” as emphasized by Hurst and Pugsley (2011).
Individuals are endowed with human capital, consisting of (1) entrepreneurial ability \((\theta_i)\) and (2) other employment specific skills \((\varepsilon_{ij})\) that are uncorrelated with entrepreneurial ability. Without loss of generality, we assume that \(\varepsilon_{Ei} = 0\) and \(\bar{\varepsilon}_j > 0\), 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

\[
H_{ji} = \theta_i^{\rho_j} \cdot e^{\delta_{ji}},
\]

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 \(\rho_U = 0\).\(^5\) We evaluate these assumptions empirically below. Thus, the effective human capital of individual \(i\) in salaried employment is increasing in (a) the person’s job-specific skills in salaried work \((\varepsilon_{Si})\), (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

\[
\ln I_{ji} = \rho_j \ln \theta_i + \varepsilon_{ji}.
\]

Individuals engaged in entrepreneurship combine entrepreneurial ability and physical capital \((K)\) to produce output \((Y)\) using a similar production function as in EJ:

\[
Y_i = H_{Ei}K_i^{\alpha}v_i = \theta_iK_i^{\alpha}v_i,
\]

---

\(^5\) This assumption 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 forms of self-employment activities, 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 self-employment and salaried work.
where \( 0 < \alpha < 1 \) and \( \nu_i \) is a lognormal disturbance, such that \( E[\nu_i] = 1 \), that reflects a independent and identically distributed productivity shock. 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 \((I_{El})\), equal

\[
I_{El} = \theta_i K_i^\alpha \nu_i - r_i K_i,
\]

(6)

where the price of output is one and the gross cost of capital \((r_i)\), i.e., 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. 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, collateral influences the cost of capital and endogenously influences the optimal capital stock and hence borrowing. Below and in Appendix A, we model the cost of capital as a function of collateral and entrepreneurial ability.

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

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

(7)

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

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

(8)

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

Notice the following 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, comparing equations (4) and (8), the returns to entrepreneurial ability in entrepreneurship are larger than the returns to entrepreneurial ability in salaried employment even when \( \rho_S = 1 \) because of the complementarity between entrepreneurial ability and physical capital.

**III.B. Selection into employment types**

Individuals select into employment types by comparing expected utility levels. For self-employment and salaried work, the log of expected utility is:

\[
\ln V_{j\ell} = \rho_j \ln \theta_i + \varepsilon_{j\ell} + \delta_{j\ell}, \text{ for } J = U \text{ or } S, \quad (10.1)
\]

and for entrepreneurship, log of expected utility is:

\[
\ln V_{E\ell} = \rho'_E \ln \theta_i + \alpha \rho'_E \ln \left( \frac{\alpha}{\bar{\gamma}} \right) + \ln (1 - \alpha) + \delta_{E\ell}. \quad (10.2)
\]

In comparing the logs of expected utilities across employment types, first consider human capital. Utility in entrepreneurship rises faster in \( \theta_i \) than utility rises in either self-employment or salaried work. This holds even when \( \rho_S = \rho_E = 1 \) because of the complementarity between entrepreneurial ability and physical capital. Furthermore, note that the log of expected utility in salaried employment (\( \ln V_{Si} \)) reflects both human capital that is specific to salaried employment (\( \varepsilon_{Si} \)) and entrepreneurial human capital (\( \rho_S \ln \theta_i \)). Ceteris paribus, therefore, while increases in \( \varepsilon_{Si} \) boost the relative utility of salaried employment, increases in \( \ln \theta_i \), even when \( \rho_S = 1 \), increase the relative utility of entrepreneurship. Second, as shown in equations (10.1 – 10.2), the cost of capital influences the utility of entrepreneurs, but not the utility of those working in other employment types. Put differently, liquidity constraints have the biggest adverse effect on those with the most entrepreneurial human capital. Finally, other human capital endowments (\( \varepsilon_{j\ell} \)) and preferences (\( \delta_{j\ell} \)) 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 \( \ln \theta_i \) as functions of the cost of capital (\( \eta_i \)), non-entrepreneurial human capital skills (\( \varepsilon_{U\ell} \) and \( \varepsilon_{Si} \)),
and preferences (\(\delta_{ui}\)). In particular, define \(ln\theta_{si}\) as the level of \(ln\theta_i\) such that below \(ln\theta_{si}\), the individual selects into self-employment and at \(ln\theta_{si}\), the individual is indifferent between self-employment and salaried work. Setting \(lnV_{ui}=lnV_{si}\), and solving for \(ln\theta_{si}\) yields:

\[
ln\theta_{si} = \frac{\delta_{ui}+(\varepsilon_{ui}\varepsilon_{si})}{\rho_{s}}.
\]  

Equation (11.1) indicates that individuals with higher \(r_i\) require more entrepreneurial ability to enter entrepreneurship than similar individuals with lower capital costs. Thus, the cost of capital shapes selection into entrepreneurship, but it does not shape selection into self-employed as \(r_i\) does not appear in (11.1). 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 select into entrepreneurship.

Figure 1 illustrates the relationship between the log utility in each employment type and \(ln\theta\). The horizontal line represents the log utility of self-employment (\(lnV_{ui}\)) and equals \(\varepsilon_{ui} + \delta_{ui}\). The dashed upward sloping line is the log utility of salaried employment (\(lnV_{si}\)),

---

6 Given the modeling assumptions, \(ln\theta_{ei} > 0\) unless \(\delta_{ei}\) is exceptionally large. We already defined the conditions under which \(ln\theta_{si} > 0\). The additional conditions under which \(ln\theta_{ei} > ln\theta_{si}\) are that (i) individuals do not have preferences for self-employment or entrepreneurship that are so strong that they dominate the increasing pecuniary returns to salaried work and entrepreneurship associated with entrepreneurial ability or that (ii) self-employment skills are not so valuable that they overwhelm the increasing returns to entrepreneurial ability in entrepreneurship.
where the slope is $\rho_s$. $lnV_{Si}$ intersects $lnV_{Ei}$ at the first cutoff level of entrepreneurial ability: $ln\theta_{Si}$. The solid upward sloping line is the log utility of entrepreneurship ($lnV_{Ei}$), where the slope is $\rho_E$ and where $lnV_{Ei}$ intersects $lnV_{Si}$ at the second cutoff level: $ln\theta_{Ei}$.

Figure 1 illustrates how human capital and liquidity constraints shape selection into different employment types. On human capital, Figure 1 indicates that entrepreneurs are positively selected on entrepreneurial ability, but the self-employed are negatively selected on $ln\theta_i$. Consistent with this prediction, Levine and Rubinstein (2017) show that entrepreneurs (as proxied by the incorporated self-employed) score higher on learning aptitude tests as youths, attain higher education levels, and perform tasks (even when they were salaried employees) demanding more creative and analytically-challenging cognitive skills and less manual skills than salaried workers. The opposite holds for the self-employed as proxied by the unincorporated self-employed). On liquidity constraints, Figure 1 indicates that $r_l$ shapes entry into entrepreneurship, but not into self-employment. In particular, increases in $r_l$ shift 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 constraints do not affect entry into self-employment.

**III.C. Testable implications and discussion**

The model yields two testable implications with respect to the human capital and liquidity puzzles. First, ceteris paribus, entrepreneurs are positively selected on salaried wages, but the self-employed are negatively selected on wages. Holding individual preferences, job-specific skills, and capital costs constant, individuals who sort into entrepreneurship have (1) greater entrepreneurial abilities and (2) access to higher paying salaried jobs than those who choose to work as salaried employees (S) and self-employed business owners (U). This is illustrated in Figure 1: 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 self-employment. Ceteris paribus, it is people with low entrepreneurial abilities and hence comparatively low-paying salaried options who choose to become self-employed. Second, entrepreneurs are negatively selected on the cost of capital, but the self-employed are not. That is, liquidity constraints matter for entrepreneurs but not for other self-employed business owners.

These two implications of the model are unique in that other models of entrepreneurship do not distinguish between entrepreneurs and other business owners and therefore do not derive predictions regarding the contrasting selection of individuals into entrepreneurship and self-employment on human capital and liquidity constraints. The model explains why aggregating these two groups and calling the combined group “entrepreneurs” can lead to mis-leading perspectives on entrepreneurship. Finally, the model’s prediction that entrepreneurs have higher salaried employment options than salaried workers (holding everything except entrepreneurial ability constant) is different from EJ, where wages are uncorrelated with entrepreneurial earnings.

III. D. Extension: Risk aversion

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

$$V_{jl} = -\exp\{\tau_l I_{jl} * \delta_{jl}\}, \quad (1')$$

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

$$Y_l = \theta_i K_l^\alpha (1 + \nu_l), \quad (5')$$

---

7 The model also predicts that the tendency for people with the best salaried job opportunities to become entrepreneurs is stronger in industries and activities with (1) higher $\rho$, such that entrepreneurial ability is more productive in salaried employment, and (2) higher $\alpha$, such that the returns to entrepreneurship are greater.
where $v_i'$ is a zero mean, normally distributed shock to productivity. Assuming that the variance of output is $\sigma_i^2 = \sigma^2 \theta K^\alpha$, i.e., the variance of aggregate output does not change if a firm is split into two or more firms, then the variance of $v_i'$ equals $\sigma_i^2 / \theta K^\alpha$.

Thus, expected utility in entrepreneurship is:

$$E\{V_{El}\} = -\exp \{-\tau_i[\theta_i K_i^\alpha - r_i K - \theta_i K_i^\alpha (\sigma^2 / 2)]\},$$

where, for simplicity, we have set $\delta_{El} = 0$. Exploiting the observation that the certainty equivalent earnings from entrepreneurship is $I_{El}' = \theta_i K_i^\alpha (1 - \tau_i (\sigma^2 / 2)) - r_i K_i$, the optimal capital stock for entrepreneur $i$ is:

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

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:

$$ln\{I_{El}'\} = \rho_E l\theta_i + \rho_E L\gamma_i + \alpha \rho_E L\alpha + Ln(1 - \alpha).$$

Furthermore, it is important to note 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 risk neutral specification hold: (1) entrepreneurs are positively selected on entrepreneurial ability ($ln\theta_i$) and on salaried wages (as salaried wages vary positively with $ln\theta_i$), but other self-employed business owners are negatively selected on salaried wages and (2) entrepreneurs are negatively selected on the cost of capital but other self-employed business owners 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 noncognitive attitudes toward risk ($\gamma_i$). This interaction suggests that effective entrepreneurial human capital is a mixture of narrowly defined
entrepreneurial ability and the noncognitive traits that allow individuals to exercises those skills. Thus, the most successful entrepreneurs might not be those with the most entrepreneurial ability, e.g., if risk tolerance (τ) and entrepreneurial ability (θ) are negatively correlated. This is akin to the combination of “smart and illicit” traits emphasized by Levine and Rubinstein (2017), where illicit capture attitudes toward breaking from the norm, undertaking novel endeavors, and investing in risky ventures. In the remainder of this paper, we focus on the risk neutral case and evaluating its predictions and show that the results are robust to controlling for other preferences and characteristics associated entrepreneurship.

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) it takes very little capital to start most U.S. businesses and (2) the aggregate group of self-employed business owners have very similar human capital traits as their salaried counterparts. Our model provides novel predictions about the impact of human capital and liquidity constraints on selection into entrepreneurship, salaried employment, and self-employment. Thus, our model motivates a reconsideration of the human capital and liquidity puzzles. In this section, we take the theoretical model from section III and derive estimable equations that will allow us to reassess the roles of human capital and liquidity constraints in shaping selection across employment types.

In moving from the model toward an estimable equation, first note from equation (4) that the log of potential wages from salaried employment (\( W_i \)) is \( W_i = \rho_s \ln \theta_i + \varepsilon_{sl} \). We use the term “potential” wages because the individual considers this wage in selecting an employment type, but will not receive this wage if the individual chooses to become self-employed or an entrepreneur. Next, let \( C_i \) equal the collateral of individual \( i \) and assume that the elasticity of \( r_l \)

---

8 Note that selection on entrepreneurial ability might vary across industries if \( \sigma^2 \) differs across industries.
with respect to collateral is negative. Incorporating these two features into equations (10.1) and (10.2) and substituting out for $\ln \theta_i$ and $\eta_i$, we can rewrite the log expected utility of individual $i$ in each employment type $J$ as:

$$lnV_{ji} = b_{jW}W_i + b_{jC}C_i + b_{jX}X_i + \mu_{ji},$$  \hspace{1cm} (12)

where the error term combines person-specific preferences for employment type $J$ ($\delta_{ji}$) and person-specific skills in employment type $J$ ($\epsilon_{ji}$). With respect to collateral, the model predicts that $b_{SC} = 0$, $b_{UC} = 0$, and $b_{EC} > 0$. With respect to potential wages ($b_{jW}$), equation (10.1) implies that $b_{SW} = 1$ and $b_{UW} = 0$. Furthermore on potential wage, the derivation of equation (12) indicates that (a) $b_{EW} > 0$ and (b) $b_{EW}$ can be greater than $b_{SW}$. Intuitively, if there is not much cross-sectional variance in $\epsilon_{si}$, then cross-sectional differences in potential salaried wages provide a good signal of differences in entrepreneurial ability that are reflected in even bigger differences in entrepreneurial earnings.

The next challenge in deriving an estimable equation is that potential wages ($W_i$) are unobservable, e.g., we do not observe actual salaried wages for those who sort into self-employment or entrepreneurship. To address this challenge, we exploit the observation that almost all individuals work as salaried workers before becoming business owners and use these early career wages to proxy for potential wages later in life. We evaluate this assumption below.

Finally, employment type choices in relative expected utility terms yields the multinomial logit equations that we estimate in the next section. Specifically, after (1) assuming that wages early in a person’s work life ($W_{Si0}$) are positively correlated with potential wages later in life (while conditioning on demographics and other cognitive and noncognitive traits), (2) defining the net log expected utility of employment type $J$ relative to salaried employment for individual $i$

---

9 For example, if $\alpha/\tau_i = \exp(\kappa C)$, then $\ln(\alpha/\tau_i) = \kappa C$. Assuming that $\kappa > 0$ implies that $b_{EC} = \kappa \alpha \rho'_e > 0$ in equation (12).

10 From Appendix B, $b_{EW} = \rho_e \rho_s \sigma^2_\delta / (\rho^2_e \sigma^2_\delta + \sigma^2_{Es})$, where $\sigma^2_\delta$ is the variance of $\ln \theta_i$, $\sigma^2_{Es}$ is the variance of job-specific skills in salaried work. Thus, if $\sigma^2_{Es}$ is sufficiently small, then $b_{EW} > b_{SW} = 1$. 
in each period as $U_{jt}$, where $U_{jt} = \ln V_{jt} - \ln V_{st}$, and (3) substituting these expressions into equation (12), we obtain the following:

$$U_{jt} = \beta_{jw} W_{sito} + \beta_{jc} C_{it} + \beta_{jx} X_{it} + \eta_{jit},$$

(13)

where $\eta_{jit}$ includes shocks to earnings in period $t$. Thus, the probability that individual $i$ prefers entrepreneurship or self-employment to salaried employment is

$$P(U_{jit} > U_{sit}) = P(\beta_{jw} W_{sito} + \beta_{jc} C_{it} + \beta_{jx} X_{it} > [\eta_{sit} - \eta_{jit}]).$$

(14)

This directly motivates an estimable multinomial logit regression:

$$\ln\left(\frac{P_{jit}}{P_{sit}}\right) = \beta_{jw} W_{sito} + \beta_{jc} C_{it} + \beta_{jx} X_{it},$$

(15)

where the dependent variable is the log-odds ratio of being an entrepreneur (self-employed) business owner rather than a salaried worker, where $P_{jit}$ stands for the probability that person $i$ is an entrepreneur ($J=E$) or self-employed ($J=U$) at time $t$, $P_{sit}$ denotes the probability that the person is a salaried worker at time $t$.

There are two reduced form parameters of interest: selection on wages ($\beta_{jw}$) and the impact of collateral on entry into entrepreneurship and self-employment ($\beta_{jc}$). The model yields unambiguous predictions about collateral: $\beta_{uc} = 0$ and $\beta_{ec} > 0$; that is, collateral does not shape preferences with respect to salaried employment or self-employment, but collateral lowers the costs of becoming an entrepreneur. With respect to wages, the model unambiguously predicts that $\beta_{uw} < 0$; increases in wages increase the utility of salaried employment relative to self-employment. The model yields ambiguous predictions with respect to $\beta_{ew}$. To the extent that wages are higher because the individual has higher salaried-specific skills, then $\beta_{ew} < 0$, as wages rise but entrepreneurial earnings do not. To the extent that wages are higher because entrepreneurial ability is higher, then $\beta_{ew} > 0$; entrepreneurial ability increases earnings more in entrepreneurship than in salaried employment.
V. EMPIRICAL RESULTS

In this section we empirically evaluate the model’s two key predictions: (1) entrepreneurs are positively selected on salaried wages and associated human capital, but other business owners are negatively selected on those factors and (2) entrepreneurs are positively selected on collateral, but not other business owners. With respect to collateral, we exploit exogenous variation in home equity values and also assess the differential impact of collateral on entry into entrepreneurship and other businesses. In particular, Section B examines selection on wages and home while, while Section C evaluates the causal impact of collateral on entry into incorporated and unincorporated self-employment. We begin, in Section A, by assessing a key ingredient in the model: that early career wages are positively associated with both future salaried wages and entrepreneurial earnings but not earnings as an unincorporated business owner.

V.A. Early career wages and future earnings

As discussed above and developed more formally in Appendix B, the model’s key assumptions imply that early career wages are positively associated with future wages and entrepreneurial earnings but not with future earnings from other forms of self-employment. The model’s assumptions also indicate that the relationship between early career wages and future entrepreneurial earnings can be even larger than the relationship with future salaried wages if (a) early salaried wages are a good indicator of entrepreneurial skills (i.e., if \( \rho \) is large) and (b) the returns to entrepreneurial skills are sufficiently large in entrepreneurship (i.e., if \( \alpha \) is large).

To evaluate these predictions, Table III provides estimates from a log linear regression of log hourly earnings in employment type \( J \) in year \( t \) (\( W_{jit} \)) on \( W_{si0} \), which equals log hourly salaried earnings when the individual was 25-29 years old (Wages (25-29))

\[
W_{jit} = a_j + y_j^W W_{si0} + y_j^X X_i + u_{jit},
\]

where we examine individuals who are older than 30. For employment types \( J=U, S, \) and \( E \), we unincorporated business owners, salaried employees, and incorporated business owners respectively. We also provide results for the aggregate group of self-employed individuals, i.e.,
combining unincorporated and incorporated businesses owners. $X_i$ is a set of control variables for individual $i$. In Panel A, these control variables include an individual’s education, experience, gender, race, cognitive and noncognitive traits (AFQT, Locus of control, Self-esteem index), as well as state and year effects. In Panel B, we also control for state-year fixed effects. We restrict the sample to individuals who are full-time, full-year workers in year $t$.

The results in Table III do not reject the modeling assumptions that yield predictions concerning $\gamma^W_j$. First, we cannot reject the hypothesis that $\gamma^W_s = 0$ as shown in columns (3) and (7). Second, columns (1) and (5) indicate that $0 < \gamma^W_s < 1$. Finally, the results in columns (4) and (8) show that $\gamma^W_e > 0$. In fact, we also find that $\gamma^W_e > \gamma^W_s$, suggesting both that entrepreneurial skills are strongly positively associated with salaried earnings and that entrepreneurial ability is a large component of entrepreneurial production.

**V.B. Selection on wages and home wealth**

We next examine the differential selection into entrepreneurship and unincorporated self-employment on early careers wages and collateral. Based on equation (15), we focus on estimating the multinomial logit regression:

$$
\ln \left( \frac{P_{jt \mid W}}{P_{jt \mid U}} \right) = \beta_w W_{S_{t0}} + \beta_c C_{it} + \beta_x X_{it} + u_{jt},
$$

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 IV, we provide the multinomial logit results on unincorporated and incorporated self-employment (columns 4-5), where we do not report results on other self-employment categories such as unpaid family and nonprofit businesses. For completeness, we also provide the results from logit regressions (columns 1-3). In these logit regressions, the dependent variable is an indicator variable that equals one if the individual is self-employed (column 1), unincorporated (column 2), or incorporated (column 3) in year $t$ respectively and zero otherwise. The key explanatory variables are *Wages* (25-29), which equals log hourly salaried earnings when the individual was 25-29 years old ($W_{S_{t0}}$), and *Home Wealth*(t-2), which equals the market value of the individual’s
home (if any) minus mortgages on the house divided by $100,000 ($C_{it}$). Although unreported, all regressions condition on Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), cognitive and non-cognitive traits (AFQT, Self-esteem, Locus of Control), gender, race, year, and state fixed effects. The logit regressions also include gender-year and race-year effects. The sample is restricted to 1987-2014 and excludes individuals who were self-employed in either $t-2$ or $t-4$ to focus on entry into business ownership. The table provides heteroskedasticity robust standard errors that are clustered at the individual level.

Consistent with the model’s predictions, the results reported in Table IV indicate positive selection into incorporated self-employment on early career wages and negative selection into unincorporated self-employment on those salaried earnings. $Wages (25-29)$ enters positively and significantly when examining selection into incorporated self-employment (columns 3 and 5) but negatively and significantly when assessing entry into unincorporated self-employment (columns 2 and 4). 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 15.5% greater than the low early-career wage worker ($1.155=\exp(0.5*0.2887)$). Similarly, the estimated coefficients indicate that the odds of the low early-career wage earner switching from salaried work into unincorporated business ownership next period are 20.5% greater than the high early-career wage worker ($1.205=\exp(0.5*0.3746)$). Table IV also highlights the pitfalls of using the aggregate group of self-employed business owners. As shown in the column (1), there is negative selection into self-employment on early career wages for the aggregate group of self-employed, which masks the differential selection into entrepreneurship and other businesses.
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-employment. That is, $Home Wealth(t-2)$ enters positively, significantly, and with an economically large coefficient when examining incorporated self-employment (columns 3 and 5) but enters with a small, insignificant coefficient when examining unincorporated self-employment (columns 2 and 4). With respect to the economic size of the estimated coefficients, consider (a) the multinomial logit regression results reported in column (5) and (b) a high-collateral and low-collateral person, where the high-collateral person has $50,000 of additional home wealth in year $t-2$ 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 8% greater than the low-collateral person ($=\exp(0.5*0.1547)$).

We next address two potential omitted variable concerns. First, there might be differences in local economic conditions that shape early career wages, collateral, and selection into different employment types. Omitting these factors from the Table IV analyses could generate spurious results. To address this concern, we include state-year fixed effects to control for all time-varying state influences. Second, time-invariant individual traits might drive the results on collateral. For example, individuals from rich families might have many characteristics that facilitate entry into entrepreneurship. If collateral is positively correlated with these other characteristics, the Table IV results might inappropriately conclude that collateral shapes entry into entrepreneurship. To address this concern, we include individual fixed effects. In this way, we focus on whether changes in collateral are associated with selection into different employment types. Here, we provide these assessments using individual fixed effects; below, we implement a strategy for identifying the impact of collateral on selection into employment types below. Of course, including individual fixed effects will essentially eliminate $Wages (25-29)$, as it varies little over
time.\(^{11}\) When conducting these analyses with individual fixed effects, we use OLS, as the multinominal logit regressions did not converge. For comparison purposes, therefore, we present all of these additional analyses using a linear probability model.

Table V shows that the results are robust to implementing these strategies. First, after controlling for state-year fixed effects (Panel A), we continue to find (1) positive selection into entrepreneurship on early career wages and negative selection into unincorporated self-employment on those wages and (2) positive selection into incorporated self-employment on collateral, but not into unincorporated self-employment. Second, when conditioning on individual fixed effects, we continue to find that positive selection into entrepreneurship on \(\text{Home Wealth}(t-2)\).

We also provide a falsification test in Table VI of the results emerging from Tables IV and V. Instead of examining \(\text{Home Wealth}(t-2)\), we examine \(\text{Home Wealth}(t+2)\). If \(\text{Home Wealth}(t-2)\) captures changes in wealth that can be used as collateral to finance entry into entrepreneurship in year \(t\), then we expect \(\text{Home Wealth}(t-2)\) to enter positively—and \(\text{Home Wealth}(t+2)\) to enter insignificantly—into the incorporated self-employment regressions when controlling for individual fixed effects. We would not expect that a change in future household wealth would influence past entry into entrepreneurship unless \(\text{Home Wealth}(t+2)\) is capturing something else about the evolving characteristics of the individual. This is what we discover in Table VI: When controlling for individual and state-year fixed effects, we find positive selection into entrepreneurship on \(\text{Home Wealth}(t-2)\) but not on \(\text{Home Wealth}(t+2)\).

\textit{V.C. The impact of home wealth on entry into business ownership}

Although the results reported in Tables IV-VI 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 and state-

\(^{11}\) As noted in Section II, there is slight time variation in Wages (25-29) when individuals are 27-30 years old.
year fixed 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), Coradin and Popov (2015), and Schmalz, Sraer, and Thesmar (2017), we use \( \text{Home Wealth}_{it-4} \times g_{(t-4,t-1)} \), which equals the net value of the home owned by individual \( i \) in year \( t-4 \) (\( \text{Home Wealth}_{it-4} \)) times the growth rate of home prices in the state in which the home is located from year \( t-4 \) to year \( t-1 \) (\( g_{(t-4,t-1)} \)). If the individual does not own a home in year \( t-4 \), \( \text{Home Wealth}_{it-4} \) 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 several advantages over previous authors. Since we have panel data, we (a) use individual fixed effects as discussed above, (b) conduct the analyses based on a person’s home wealth in \( t-4 \) and changes in the state’s home values through period \( t-1 \) in examining self-sorting into business ownership in period \( t \), and (c) implement several falsification tests to assess whether future home wealth or future changes in a state’s home prices shape entry into business ownership.

Thus, we estimate the following, benchmark multinomial logit model and report the results in Table VII:

\[
\ln \left( \frac{p_{j|it}}{p_{st|it}} \right) = \beta_{jW} Wages(25 - 29) + \beta_{jC} \text{Home Wealth}_{inst-4} \times g_{(t-4,t-1)} + \beta'_{jX} X_{it} + u_{jit}, \tag{17}
\]

where \( X \) includes \( \text{Home Wealth}_{it-4} \) and \( g_{(t-4,t-1)} \), as well as 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), gender, race, year, and state fixed effects. The dependent variable is the log-odds ratio of entry into either incorporated or unincorporated self-employment relative to not switching into self-employment. 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 \( \text{Home Wealth}_{it-4} \times g_{(t-4,t-1)} \) predicts \( \text{Home Wealth}_{it} \) after controlling for \( \text{Home Wealth}_{it-4}, g_{(t-4,t-1)}, \) \( \text{Wages}(25-29), \) and \( X. \) We conduct these analyses using OLS in columns (1) and (2) of Table VII, where the column (2) regression includes individual fixed effects. As shown, \( \text{Home Wealth}_{it-4} \times g_{(t-4,t-1)} \) 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 overall business ownership, i.e., the aggregate group of incorporated and unincorporated 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), \( \text{Home Wealth}_{it-4} \times g_{(t-4,t-1)} \) does not help account for entry into 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 in Table VII, we discover that collateral impacts entry into entrepreneurship but not into unincorporated business ownership. As shown in columns (5) and (6), \( \text{Home Wealth}_{it-4} \times g_{(t-4,t-1)} \) 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 \( (1.04=\exp(0.25*0.1583)). \)

One might be concerned that (a) housing price growth in a state is 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, 

\[ Home\ Wealth_{t-4} \times g(t-4, t-1) \]

might proxy for the interactive impact of (non-home wealth) individual traits and changes in overall economic conditions on entry into entrepreneurship, such 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 control for shocks to overall economic conditions and the interaction with housing wealth. To measures shocks to overall economic conditions, we use the change in the state unemployment rate (individual i’s state) between \( t-1 \) and \( t \) (\( \Delta Unemployment \)), and we also include its interaction with housing wealth in period \( t-4 \) (\( Home\ Wealth_{t-4} \times \Delta Unemployment \)). As shown in columns (7) and (8), the results do not change. Neither \( \Delta Unemployment \) nor \( Home\ Wealth_{t-4} \times \Delta Unemployment \) enters significantly. Moreover, the estimated coefficient on \( Home\ Wealth_{t-4} \times g(t+1, t+4) \) remains positive, significant, and virtually unchanged in the incorporated self-employment equation (column 8). These robustness tests are consistent with the view that collateral positively impacts the odds of switching into entrepreneurship but is not significantly linked with the odds of becoming an unincorporated business owner.

We also considered a falsification test to address the concern that \( Home\ Wealth(t-4) \times g(t-4, t-1) \) is capturing something else about an individual besides a shock to home wealth. If \( Home\ Wealth(t-4) \times g(t-4, t-1) \) captures only shocks to a person’s home wealth between year \( t-4 \) and \( t-1 \) that can be used as collateral to finance entry into entrepreneurship in year \( t \), then we expect both that (a) \( Home\ Wealth(t-4) \times g(t-4, t-1) \) will positively influence selection into incorporated self-employment analyses in year \( t \) and (b) \( Home\ Wealth(t-4) \times g(t+1, t+4) \) 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. This is what we discover in Table VIII: \( Home\ Wealth(t-4) \times g(t+1, t+4) \) 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: Entrepreneurial ability, capacity, and liquidity

Levine and Rubinstein (2017) discover that the combination of “smart” ($AFQT$) and “illicit” tendencies (aggressive, risk-taking, disruptive, "break-the-rules" behaviors) helps explain who becomes a successful entrepreneur. This is consistent with the view that the mixture of strong analytical skills and break-from-the-norm, risk-tolerant preferences shapes entry into entrepreneurship and success as an entrepreneur.

Our theory provides a possible explanation for these findings and offers additional testable implications. The model developed above demonstrates that effective entrepreneurial human capital reflects both entrepreneurial ability and risk tolerance and shows that effective entrepreneurial human capital is positively associated with selection into entrepreneurship but not into other forms of business ownership. To the extent that $AFQT$ and Illicit are positively associated with entrepreneurial ability and risk tolerance respectively, then the model predicts that the combination of “smart and illicit” traits should be positively associated with successful entrepreneurship when holding other things constant, including the independent influences of $AFQT$ and Illicit.

Thus, we next introduce smart and illicit into our analysis. Besides controlling for early career wages, Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of several cognitive and non-cognitive traits ($AFQT$, Rosenberg self-esteem, Rotter Locus of Control), $Home\ Wealth_{it-4}$, $g_{(t-4,t-1)}$, and gender, race, year, and state fixed effects, we now also control for $Smart\ &\ Illicit$ and $Illicit$. As defined above and in the Data Appendix, (a) $Illicit$ measures the degree to which an individual reported in the 1980 NSLY79 survey engaging in a range of illicit activities before reaching prime working-age, (b) $Illicit$ has a mean of zero and standard deviation of one, and (c) $AFQT$ measures cognitive abilities and is calculated as a percentile of the NLSY79 survey, with
a median value of 50. From these variables, we create 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.

As shown in Table IX, there are two key findings associated with introducing Smart & Illicit into the analyses. First, we confirm the results from Table VII: (a) shocks to collateral (Home Wealth_{it-4} * g_{(t-4,t-1)}) are positively associated with selection into incorporated self-employment (columns 4 and 6) but not into unincorporated self-employment (columns 3 and 5) or the aggregate group of self-employed (columns 1 and 2), and these results hold when controlling for shocks to economic conditions in a state (ΔUnemployment) and the interaction between those shocks and an individual’s housing wealth (Home Wealth_{it-4} * ΔUnemployment); and (b) early career wages are positively associated with selection into incorporated self-employment (columns 4 and 6) and negatively associated with selection into unincorporated self-employment (columns 3 and 5). Second, we find that Smart & Illicit is positively associated with selection into incorporated self-employment and negatively associated with entry into unincorporated self-employment, confirming the findings in Levine and Rubinstein (2017) while further conditioning on the shocks to the individual’s collateral constraints.

VI. CONCLUSIONS

Although theories stress that human capital and liquidity constraints influence entry into entrepreneurship, the data indicate that (1) salaried employees and self-employed business owners have remarkably similar levels of human capital and (2) most of the self-employed start with little or no capital. To address the gap between theory and evidence, we first offer a new three-sector Roy model of selection into entrepreneurship and other businesses on human capital and liquidity constraints. We then evaluate empirically the model’s predictions.

The distinction between entrepreneurs and other self-employed individuals is a key starting point in both the theoretical model and the empirical evaluation. Some of the self-
employed are entrepreneurs, but many are one-person businesses that perform routine, manual tasks. In the theoretical model, therefore, we distinguish between entrepreneurs and other business owners and examine selection into each type of self-employment. In the empirical evaluation, we use the incorporated self-employed as a proxy for entrepreneurs and the unincorporated as a proxy for other self-employed individuals. As shown by LR, the incorporated and their businesses perform activities requiring creativity, complex problem-solving, and the ability to persuade, motivate, and manage others. We view these human capital skills as consistent with core conceptions of entrepreneurship. In contrast, the unincorporated and their businesses engage in activities demanding manual dexterity.

The theoretical model yields distinct predictions about selection into entrepreneurship, other businesses, and salaried jobs on human capital and liquidity. Specifically, it predicts that entrepreneurs are positively selected on wages and associated human capital skills; other business owners are negatively selected on these same features; entrepreneurs are positively selected on collateral; but entry into other businesses is unrelated to liquidity. Thus, the model suggests that existing puzzles concerning entrepreneurship, human capital, and liquidity constraints might reflect the failure to distinguish between entrepreneurs and other business owners.

Consistent with the theoretical model, we discover that the incorporated are positively selected on salaried wages, the unincorporated are negatively selected on wages, collateral exerts a large, positive impact on entry into incorporated self-employment, but collateral does not influence entry into unincorporated business ownership. Thus, the results highlight the empirical shortcoming of using the aggregate group of business owners to assess selection into entrepreneurship as human capital and liquidity constraints shape entry into entrepreneurship very differently from entry into unincorporated self-employment.
REFERENCES


APPENDIX A: MODEL EXTENSIONS

This Appendix explores the implications of relaxing the assumptions of (a) zero correlation between the cost of capital and collateral, (b) zero correlation between the cost of capital and entrepreneurial ability, and (c) zero correlation between collateral and entrepreneurial ability.

To examine the implications of allowing the cost of capital to vary with collateral and entrepreneurial ability, let

$$\ln(r_i) = \kappa_{\theta} \ln(\theta_i) + \kappa_C \ln(C_i) + \omega_{ri},$$  \hspace{1cm} (A1)

where $\kappa_{\theta}$ is the elasticity of the cost of capital with respect to entrepreneurial ability ($\theta_i$), $\kappa_C$ is the elasticity of the cost of capital with respect to collateral ($C_i$), and $\omega_{ri}$ is that part of the log cost of capital that is uncorrelated with an individual’s entrepreneurial ability or collateral. For example, $\omega_{ri}$ could reflect national policies and economic conditions that influence the cost of capital. If the cost of capital is not positively associated with an individual’s collateral and entrepreneurial ability, then $\kappa_{\theta} \leq 0, \kappa_C \leq 0$.

Similarly, to allow entrepreneurial ability to vary with collateral, let

$$\ln(C_i) = \kappa_{\theta} \ln(\theta_i) + \omega_{ci},$$  \hspace{1cm} (A2)

where $\kappa_{\theta}$ is the elasticity of collateral with respect to entrepreneurial ability ($\theta_i$) and $\omega_{ci}$ is that part of collateral that is uncorrelated with an individual’s entrepreneurial ability. We assume that collateral is not negatively associated with entrepreneurial ability, i.e., $\kappa_{\theta} \geq 0$. A positive correlation could arise if family wealth both increases an individual’s collateral and provides for greater educational and social connections that increase the person’s productivity as an entrepreneur.

Substituting (A1) and (A2) into (11) yields an augmented expression for the log utility of engaging in entrepreneurship $lnV_{Ei}'$.
\[ V'_{Ei} = (1 - \alpha(\kappa_{r\theta} + \kappa_{rC}\kappa_{c\theta}))\rho'_E\ln(\theta_i) + \alpha\rho'_E(\omega_{ri} + \kappa_{rC}\omega_{Ci}) + \lambda, \]  
\[ (A3) \]

where \( \lambda = \alpha\rho'_E\ln(\alpha) + \ln(1 - \alpha) + \delta_{Ei}. \)

Furthermore, the new level of entrepreneurial ability such that above this level (\( \ln\theta'_{Ei} \)) individuals select into entrepreneurship and at this level, individuals are indifferent between salaried employment and entrepreneurship becomes:

\[ \ln\theta'_{Ei} = \frac{\alpha\rho'_E(\omega_{ri} + \kappa_{rC}\omega_{Ci}) - \lambda + \delta_{s_i}}{\rho'_E - \rho_s}. \]  
\[ (A4) \]

There are several noteworthy points. First, equation (A3) indicates that the elasticity of entrepreneurial earnings with respect to entrepreneurial ability will be greater when the cost of capital is more sensitive to entrepreneurial ability, i.e., when \( \kappa_{r\theta} \) is more negative. Thus, allowing for a negative correlation between the cost of capital and entrepreneurial ability magnifies the effects discussed in the text. Second, collateral has a direct, positive effect on entrepreneurial utility by lowering the cost of capital. In particular, greater collateral (\( \omega_{Ci} \)) (1) lowers the log cost of capital by \(-\kappa_{rC}\) (equation A1 and A2) and (2) increases entrepreneurial utility by \(-\alpha\rho'_E(\kappa_{rC})\) (equation A3). Third, if collateral is positively associated with entrepreneurial ability (\( \kappa_{C\theta} > 0 \)), then the elasticity of entrepreneurial earnings with respect to entrepreneurial ability will be greater when the cost of capital falls with collateral, i.e., when \( \kappa_{rC} \) is more negative. This magnifies the effects of entrepreneurial ability on entrepreneurial earnings discussed in the text. By relaxing the assumptions that \( \kappa_{r\theta} = \kappa_{rC} = \kappa_{C\theta} = 0 \), this appendix suggests that the relationships between collateral, entrepreneurial ability, and the cost of capital tend to magnify the effects discussed in the text.
APPENDIX B: ASSESSING KEY MODELING ASSUMPTIONS

This Appendix evaluates the assumptions: $\rho'_E > \rho_S > \rho_U \geq 0$. We show how to conduct this evaluation using salaried wages early in a person’s career, which is readily observable, rather than entrepreneurial ability, which is not. In implementing this assessment, we also provide evidence on the validity of using salaried wages early in a person’s career as a proxy for entrepreneurial ability in subsequent analyses.

Although we evaluate the two key implications of the model regarding the differential selection on wages and liquidity constraints into entrepreneurship and other forms of self-employment in Section IV of the paper, it is also valuable to examine the model’s key assumption: $\rho'_E > \rho_S > \rho_U \geq 0$. That is, the pecuniary returns to log entrepreneurial ability are greater in entrepreneurship ($E$) than in salaried work ($S$), and the returns to $\ln \theta_i$ are greater in salaried work than in self-employment ($U$).

We can provide empirical evidence on whether the data refute these key assumptions even though $\ln \theta_i$ is unobservable. We do this by examining the relationship between salaried wages early in a person’s career—and before they potentially select into entrepreneurship or self-employment—and earnings later in the person’s career in employment types $E$, $S$, and $U$. This assessment not only sheds empirical light on key modeling assumptions; it also provides evidence on the validity of using salaried wages early in a person’s career as a proxy for entrepreneurial ability in subsequent analyses.

More specifically, we express log earnings in employment type $J$ in period $t=1$ ($w_{j11}$) as a function of salaried wages early in a person’s career, i.e., in $t=0$ ($w_{S10}$):

$$w_{j11} = a_j + \gamma_j w_{S10} + u_{j11}, \quad (B.1)$$

where $u_{j11}$ ($u_{j10}$) is an i.i.d., zero mean shock to $w_{j11}$ ($w_{j10}$). From a regression, therefore, the estimated value of $\gamma_j$ is

$$\gamma'_j = \frac{\text{cov}(w_{j11}, w_{S10})}{\text{var}(w_{S10})} \quad (B.2)$$
From the model, we can derive expressions for $\gamma_j'$. From equation (4),

$$var(w_{Si0}) = \rho_\theta^2 \sigma_\theta^2 + \sigma_{\varepsilon}^2 + \sigma_{uSi0}^2,$$

where $\sigma_\theta^2$ is the variance of $Ln\theta_j$, $\sigma_{\varepsilon}^2$ is the variance of job-specific skills in salaried work ($\varepsilon_{Si}$), and $\sigma_{uSi0}^2$ is the variance of $u_{j10}$. From equations (4) and (8), the values of $cov(w_{j11}, w_{Si0})$ for $J=U, S,$ and $E$ are given by:

$$cov(w_{U11}, w_{Si0}) = \rho_U \rho_S \sigma_\theta^2 + \sigma_{eUS},$$

(B.4.1)

$$cov(w_{S11}, w_{Si0}) = \rho_S^2 \sigma_\theta^2 + \sigma_{\varepsilon}^2,$$

(B.4.2)

$$cov(w_{E11}, w_{Si0}) = \rho_E^2 \rho_S \sigma_\theta^2,$$

(B.4.3)

where $\sigma_{eUS}$ is the covariance between job-specific skills in self-employment and salaried work.

The model’s key assumptions then predict:

(1) $\gamma_U' \approx 0$, under the maintained assumption that $\rho_U = 0$ and the additional assumption that the covariance between job-specific skills in self-employment and salaried work is about zero ($\sigma_{eUS} \approx 0$);

(2) $0 < \gamma_S' < 1$, under the additional assumption that $\sigma_{uSi0}^2 > 0$; and

(3) $\gamma_E' > 0$, under the assumption that $\rho_S > 0$. In particular, the assumption that $\rho_S > 0$ implies that salaried wages provide a positive signal about entrepreneurial ability. Furthermore, the model allows for the possibility that $\gamma_E' > \gamma_U'$, which will occur if salaried wages are a good proxy for entrepreneurial skills ($\rho_S$ is large) and the returns to entrepreneurial ability in entrepreneurship are sufficiently large ($\alpha$ is large), and the variance of job-specific salaried skills is not too big ($\sigma_{\varepsilon}^2$).

To evaluate these predictions, Table III provides estimates of equation (B.1),

$$w_{j1t} = a_j + \gamma_j w_{Si0} + u_{j1t}.$$

To measure $w_{Si0}$, we use the log of hourly salaried earnings when the individual is between 25 and 29 years old. We measure $w_{j1t}$ as the log of earnings in employment type $J$ in year $t$ for

---

12 As shown in Part A of this Appendix, we can allow entrepreneurial ability to be correlated with the cost of capital. In this case, $cov(w_{E11}, w_{Si0}) = \rho_E^2 \rho_S \sigma_\theta^2 (1 + \alpha \kappa_\theta)$, where $\kappa_\theta$ is the elasticity of the cost of capital with respect to entrepreneurial ability. To the extent that entrepreneurial ability and the cost of capital are negatively correlated, $cov(w_{E11}, w_{Si0})$ becomes larger.
individuals older than 30. For employment types \(J=U, S, \) and \(E\), we define \(U\) as individuals who are unincorporated business owners, \(S\) as individuals who are salaried employees, and \(E\) as individuals who are incorporated business owners. We also provide results for the aggregate group of self-employed individuals, i.e., combining \(U\) and \(E\). In Panel A, we condition on an individual’s education, experience, gender, race, cognitive and noncognitive traits (AFQT, Rotter locus of control, Rosenberg self-esteem index), as well as state and year effects. In Panel B, we also control for state-year fixed effects. We restrict the sample to individuals who are full-time, full-year workers in year \(t\).

As shown in Table 3, the data do not reject the model’s predictions concerning \(\gamma_j'\). First, we cannot reject the hypothesis that \(\gamma_U' = 0\) as shown in columns (3) and (7). Second, columns (1) and (5) indicate that \(0 < \gamma_S' < 1\). Finally, the results in columns (4) and (8) show that \(\gamma_E' > 0\). In fact, we also find that \(\gamma_E' > \gamma_S'\), suggesting both that entrepreneurial skills are strongly positively associated with salaried earnings \((\rho_S > 0)\) and entrepreneurial ability is a large component of entrepreneurial production, i.e., \(\alpha\) is large.
<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 (25-29)</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>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of starting capital</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Family saving</td>
<td>74%</td>
<td>72%</td>
<td>81%</td>
<td></td>
</tr>
<tr>
<td>None needed</td>
<td>17%</td>
<td>21%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>External</td>
<td>11%</td>
<td>6%</td>
<td>24%</td>
<td></td>
</tr>
<tr>
<td>Loan: business/etc</td>
<td>4%</td>
<td>2%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Home equity</td>
<td>3%</td>
<td>1%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Loan: bank/finance</td>
<td>5%</td>
<td>3%</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Credit card</td>
<td>3%</td>
<td>4%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td>1%</td>
<td>1%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3%</td>
<td>3%</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. Table A1 provides variable definitions.
<table>
<thead>
<tr>
<th></th>
<th>Legal Form of Business: Incorporated</th>
<th>Starting Capital</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>Probit (2)</td>
</tr>
<tr>
<td>Family saving</td>
<td>0.006</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>Home equity</td>
<td>0.277**</td>
<td>0.753**</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.328)</td>
</tr>
<tr>
<td>Credit card</td>
<td>-0.024</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.278)</td>
</tr>
<tr>
<td>Loan: business/family/friend</td>
<td>0.216**</td>
<td>0.572**</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.274)</td>
</tr>
<tr>
<td>Government</td>
<td>0.063</td>
<td>0.200</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.436)</td>
</tr>
<tr>
<td>Loan: bank/finance</td>
<td>0.226***</td>
<td>0.607***</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.205)</td>
</tr>
<tr>
<td>Other</td>
<td>0.292***</td>
<td>0.796***</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.221)</td>
</tr>
<tr>
<td>None needed</td>
<td>-0.155**</td>
<td>-0.474**</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.203)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.329***</td>
<td>-0.467***</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.177)</td>
</tr>
<tr>
<td>Observations</td>
<td>1105</td>
<td>1105</td>
</tr>
<tr>
<td>R-square / Pseudo R-square</td>
<td>0.074</td>
<td>0.0574</td>
</tr>
</tbody>
</table>

Note: This table provides regression results on the relationship between the sources of starting capital and (1) the legal form of the business: Incorporated (columns 1-3) and (2) Starting Capital (columns 4-5). The explanatory variables in all five regressions are indicator variables of whether the individual reports that each particular source of capital was used to establish or acquire the business. We use OLS and Probit estimators in columns (1) and (2) respectively, where the dependent variable equals one if the legal form of the business is incorporated and zero for other legal forms. Data on the legal form of the business are from the 2010 and 2012 business ownership part of the NLSY79 survey. Column (3) provides multinomial logit estimates, where the unreported categories for the business's legal form are unincorporated and other legal forms. For the regressions in which the dependent variable is the amount of starting capital, we report OLS (column 4) and median regressions (column 5). Table A1 provides variable definitions. The sample is based on only the 2010 and 2012 business ownership part of the NLSY79 survey. Heteroskedasticity robust standard errors are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.
### TABLE III
EARNINGS BY EMPLOYMENT TYPE AND EARLY SALARIED WAGES

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

<table>
<thead>
<tr>
<th></th>
<th>Salaried (1)</th>
<th>Log Hourly Earnings (31+)</th>
<th>Self-Employed (2)</th>
<th>Unincorporated (3)</th>
<th>Incorporated (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>0.510***</td>
<td>0.2367*</td>
<td>0.067</td>
<td>0.936***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.1369)</td>
<td>(0.147)</td>
<td>(0.322)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>41015</td>
<td>3488</td>
<td>2568</td>
<td>920</td>
<td></td>
</tr>
<tr>
<td>R-square</td>
<td>0.101</td>
<td>0.0732</td>
<td>0.102</td>
<td>0.104</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Earnings vs. early salary wages, standard controls and state-year effects

<table>
<thead>
<tr>
<th></th>
<th>Salaried (5)</th>
<th>Log Hourly Earnings (31+)</th>
<th>Self-Employed (6)</th>
<th>Unincorporated (7)</th>
<th>Incorporated (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>0.5105***</td>
<td>0.2198</td>
<td>0.0275</td>
<td>1.0058**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0364)</td>
<td>(0.1480)</td>
<td>(0.1675)</td>
<td>(0.4282)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>41015</td>
<td>3488</td>
<td>2568</td>
<td>920</td>
<td></td>
</tr>
<tr>
<td>R-square</td>
<td>0.1138</td>
<td>0.2039</td>
<td>0.2698</td>
<td>0.3631</td>
<td></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. Table A1 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>Logit</th>
<th>Multinomial Logit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self Employed</td>
<td>Unincorporated</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>-0.2004***</td>
<td>-0.2914***</td>
</tr>
<tr>
<td></td>
<td>(0.0537)</td>
<td>(0.0548)</td>
</tr>
<tr>
<td>Home Wealth (t-2)</td>
<td>0.0656***</td>
<td>0.0142</td>
</tr>
<tr>
<td></td>
<td>(0.0146)</td>
<td>(0.0228)</td>
</tr>
<tr>
<td>Observations</td>
<td>105043</td>
<td>104999</td>
</tr>
</tbody>
</table>

Notes: This table reports logit (columns 1-3) and multinomial logit (columns 4-5) analyses of selection into different employment types in year t on both the individual's average log wages as a salaried employee during the ages of 25 through 29 (Wages (25-29)) and Home Wealth (t-2) two year earlier. In columns (1-3), the dependent variable is an indicator variable of whether the individual is self-employed, unincorporated self-employed, or incorporated self-employed in year t respectively. Columns (4-5) 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, race, year, and state fixed effects. The logit regressions also include gender-year and race-year effects. The sample is restricted to sample years 1987-2014 and excludes individuals who were self-employed in either t-2 or t-4. Table A1 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 HOME WEALTH: OLS AND FIXED EFFECTS**

<table>
<thead>
<tr>
<th></th>
<th>Panel A: OLS</th>
<th>Panel B: OLS with Individual Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self Employed</td>
<td>Unincorporated</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>-0.0065***</td>
<td>-0.0067***</td>
</tr>
<tr>
<td></td>
<td>(0.0018)</td>
<td>(0.0018)</td>
</tr>
<tr>
<td>Home Wealth (t-2)</td>
<td>0.0025***</td>
<td>0.0026***</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>State-year FE</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>105080</td>
<td>105080</td>
</tr>
<tr>
<td>R-square</td>
<td>0.0072</td>
<td>0.0184</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>-0.0005</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
<td>(0.0031)</td>
</tr>
<tr>
<td>Home Wealth (t-2)</td>
<td>0.0012</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>State-year FE</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>105080</td>
<td>105080</td>
</tr>
<tr>
<td>R-square</td>
<td>0.2376</td>
<td>0.2469</td>
</tr>
</tbody>
</table>

Notes: This table reports OLS analyses of selection into different employment types in year t on average log wages as a salaried employee during the ages of 25 through 29 (Wages (25-29)) and Home Wealth (t-2). The dependent variable is a one-zero indicator variable of whether the individual is self-employed, unincorporated self-employed, or incorporated self-employed in year t respectively. 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, some regressions include state-year fixed effects. In Panel B, the regressions also include individual fixed effects. The sample is restricted to sample years 1987-2014 and excludes individuals who were self-employed in either t-2 or t-4. Table A1 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 VI
SELECTION ON WAGES AND HOME WEALTH: FALSIFICATION TEST

Panel A: OLS

<table>
<thead>
<tr>
<th></th>
<th>Self Employed (1)</th>
<th>Self Employed (2)</th>
<th>Unincorporated (3)</th>
<th>Unincorporated (4)</th>
<th>Incorporated (5)</th>
<th>Incorporated (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>-0.0066***</td>
<td>-0.0068***</td>
<td>-0.0082***</td>
<td>-0.0084***</td>
<td>0.0016**</td>
<td>0.0016**</td>
</tr>
<tr>
<td></td>
<td>(0.0019)</td>
<td>(0.0019)</td>
<td>(0.0017)</td>
<td>(0.0017)</td>
<td>(0.0008)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Home Wealth(t-2)</td>
<td>0.0025***</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0006</td>
<td>0.0023***</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td></td>
<td>(0.0006)</td>
<td></td>
<td>(0.0005)</td>
<td></td>
</tr>
<tr>
<td>Home Wealth(t+2)</td>
<td>0.0027***</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0018***</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td></td>
<td>(0.0006)</td>
<td></td>
<td>(0.0005)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
</tr>
<tr>
<td>R-square</td>
<td>0.0180</td>
<td>0.0182</td>
<td>0.0192</td>
<td>0.0193</td>
<td>0.0156</td>
<td>0.0155</td>
</tr>
</tbody>
</table>

Panel B: OLS with Individual Fixed Effects

<table>
<thead>
<tr>
<th></th>
<th>Self Employed (7)</th>
<th>Self Employed (8)</th>
<th>Unincorporated (9)</th>
<th>Unincorporated (10)</th>
<th>Incorporated (11)</th>
<th>Incorporated (12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wages (25-29)</td>
<td>-0.0010</td>
<td>-0.0011</td>
<td>-0.0004</td>
<td>-0.0004</td>
<td>-0.0007</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0030)</td>
<td>(0.0028)</td>
<td>(0.0028)</td>
<td>(0.0014)</td>
<td>(0.0014)</td>
</tr>
<tr>
<td>Home Wealth(t-2)</td>
<td>0.0007</td>
<td>-0.0005</td>
<td>-0.0005</td>
<td>-0.0005</td>
<td>0.0012**</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td></td>
<td>(0.0007)</td>
<td></td>
<td>(0.0006)</td>
<td></td>
</tr>
<tr>
<td>Home Wealth(t+2)</td>
<td>0.0005</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0004</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td></td>
<td>(0.0007)</td>
<td></td>
<td>(0.0005)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
<td>99773</td>
</tr>
<tr>
<td>R-square</td>
<td>0.2574</td>
<td>0.2574</td>
<td>0.2525</td>
<td>0.2525</td>
<td>0.2459</td>
<td>0.2457</td>
</tr>
</tbody>
</table>

Notes: This table reports OLS analyses of selection into different employment types in year t on average log wages as a salaried employee during the ages of 25 through 29 (Wages (25-29)) and Home Wealth (t-2), as well as Home Wealth (t+2) as a falsification test. The dependent variable is a one-zero indicator variable of whether the individual is self-employed, unincorporated self-employed, or incorporated self-employed in year t as indicated by the column headers. 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-year fixed effects. In Panel B, the regressions also include individual fixed effects. The sample is restricted to sample years 1987-2014 and excludes individuals who were self-employed in either t-2 or t-4. Table A1 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 VII
SELECTION ON WAGES AND SHOCKS TO HOME WEALTH

<table>
<thead>
<tr>
<th></th>
<th>Home Wealth(t)</th>
<th>Self Employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(OLS) (FE)</td>
<td>(Logit) (Logit)</td>
<td>(Multinomial Logit)</td>
<td>(Multinomial Logit)</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>0.1065*** 0.0264</td>
<td>-0.1824*** -0.1817***</td>
<td>-0.3642*** 0.3132**</td>
<td>-0.3639*** 0.3139***</td>
</tr>
<tr>
<td></td>
<td>(0.0138) (0.0170)</td>
<td>(0.0584) (0.0585)</td>
<td>(0.0646) (0.1503)</td>
<td>(0.0645) (0.1502)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t-4, t-1)</td>
<td>0.6266*** 0.4883***</td>
<td>0.0002 0.0045</td>
<td>-0.0701 0.1557**</td>
<td>-0.0725 0.1583**</td>
</tr>
<tr>
<td></td>
<td>(0.0747) (0.0677)</td>
<td>(0.0482) (0.0489)</td>
<td>(0.0787) (0.0756)</td>
<td>(0.0795) (0.0767)</td>
</tr>
<tr>
<td>g(t-4, t-1)</td>
<td>0.3739*** 0.3718***</td>
<td>0.2353 0.2007</td>
<td>0.4291* -0.8914*</td>
<td>0.3828* -0.8468</td>
</tr>
<tr>
<td></td>
<td>(0.0585) (0.0519)</td>
<td>(0.2056) (0.2072)</td>
<td>(0.2228) (0.5297)</td>
<td>(0.2240) (0.5287)</td>
</tr>
<tr>
<td>Home Wealth (t-4)</td>
<td>0.7577*** 0.4993***</td>
<td>0.0655*** 0.0665***</td>
<td>0.0194 0.1665***</td>
<td>0.0229 0.1662***</td>
</tr>
<tr>
<td></td>
<td>(0.0246) (0.0217)</td>
<td>(0.0163) (0.0163)</td>
<td>(0.0278) (0.0224)</td>
<td>(0.0270) (0.0233)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*ΔUnemployment</td>
<td>-1.9228 1.2531</td>
<td>-1.6953 -1.9323</td>
<td>-2.0827 1.7655</td>
<td>-2.0827 1.7655</td>
</tr>
<tr>
<td></td>
<td>(1.2531) (2.9323)</td>
<td>(1.9323) (1.7655)</td>
<td>(2.0827) (1.7655)</td>
<td>(2.0827) (1.7655)</td>
</tr>
<tr>
<td>ΔUnemployment</td>
<td>4.2480</td>
<td>7.8769</td>
<td>-14.2894</td>
<td>12.1782</td>
</tr>
<tr>
<td></td>
<td>(5.0721)</td>
<td>(5.5167)</td>
<td>(12.1782)</td>
<td>(12.1782)</td>
</tr>
<tr>
<td>Observations</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
</tr>
</tbody>
</table>

Notes: This table reports (1) logit (columns 3-4) and multinomial logit (columns 5-8) analyses of selection into different self-employment, unincorporated self-employment, and incorporated self-employment on pre-determined salaried wages (Wages (25-29)) and a Bartik instrument for changes in home wealth (Home Wealth(t-4)*g(t-4, t-1)) and (b) OLS regressions of Home Wealth (t) on the same explanatory variables (columns 1-2). In columns (1-2), the dependent variable is the individual's home wealth in year t. In columns (3-4), the dependent variable is a one-zero indicator variable of whether the individual is self-employed in year t. For the dependent variable in columns (5-6), we use one-zero indicator variables of employment type, where the reported categories are unincorporated and incorporated, and the unreported categories are salaried, unpaid family, and other business ownership. For the Bartik instrument, 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 and Home Wealth(t-4) is the individual's net home wealth in year t-4. As indicated, some regressions control for the change in the state unemployment rate between t-1 and t (ΔUnemployment) and its interaction with Home Wealth(t-4). 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, race, year, and state fixed effects. The OLS and logit regressions also include gender-year and race-year effects. Regression (2) includes individual fixed effects (FE). The sample is restricted to sample years 1989-2014 and excludes individuals who were self-employed in either t-2 or t-4. Table A1 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 VIII
SELECTION ON WAGES AND SHOCKS TO HOME WEALTH: FALSIFICATION TEST

<table>
<thead>
<tr>
<th></th>
<th>Unincorporated (Multinomial Logit)</th>
<th>Incorporated (Multinomial Logit)</th>
<th>Unincorporated (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>-0.3639***</td>
<td>0.3139**</td>
<td>-0.3644***</td>
<td>0.3151**</td>
</tr>
<tr>
<td></td>
<td>(0.0645)</td>
<td>(0.1502)</td>
<td>(0.0646)</td>
<td>(0.1509)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t-4, t-1))</td>
<td>-0.0725</td>
<td>0.1583**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0795)</td>
<td>(0.0767)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g(t-4, t-1))</td>
<td>0.3828*</td>
<td>-0.8468</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2240)</td>
<td>(0.5287)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home Wealth (t-4)</td>
<td>0.0229</td>
<td>0.1662***</td>
<td>0.0085</td>
<td>0.1731***</td>
</tr>
<tr>
<td></td>
<td>(0.0270)</td>
<td>(0.0233)</td>
<td>(0.0302)</td>
<td>(0.0241)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t+1, t+4))</td>
<td></td>
<td></td>
<td>0.1393</td>
<td>0.0412</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.1073)</td>
<td>(0.1303)</td>
</tr>
<tr>
<td>g(t+1, t+4))</td>
<td></td>
<td>-0.0432</td>
<td>1.2652**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.2724)</td>
<td>(0.5434)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
</tr>
</tbody>
</table>

Notes: This table reports multinomial logit analyses of selection into unincorporated and incorporated self-employment on pre-determined salaried wages (Wages (25-29)) and a Bartik instrument for changes in home wealth (Home Wealth(t-4)*g(t-4, t-1)) as well as a falsification test. In columns (1-2), the dependent variable is the individual's home wealth in year t. For the dependent variable, we use one-zero indicator variables of employment type, where the reported categories are unincorporated and incorporated, and the unreported categories are salaried, unpaid family, and other business ownership. For the Bartik instrument, 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 and Home Wealth(t-4) is the individual's net home wealth in year t-4. For the falsification test, we include Home Wealth(t-4)*g(t+1, t+4)), where g(t+1, t+4)) is the growth rate in state housing prices between t+1 and t+4. 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), the change in the state unemployment rate between t-1 and t (ΔUnemployment) and its interaction with Home Wealth(t-4), as well as gender, race, year, and state fixed effects. The sample is restricted to sample years 1989-2014 and excludes individuals who were self-employed in either t-2 or t-4. Table A1 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>Self Employed</th>
<th>Unincorporated</th>
<th>Incorporated</th>
<th>Unincorporated</th>
<th>Incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Logit) (1)</td>
<td>(Multinomial Logit) (3)</td>
<td>(Logit) (2)</td>
<td>(Multinomial Logit) (4)</td>
<td>(Logit) (5)</td>
</tr>
<tr>
<td>Wages (25-29)</td>
<td>-0.1825***</td>
<td>-0.3637***</td>
<td>0.3143**</td>
<td>-0.3634***</td>
<td>0.3150**</td>
</tr>
<tr>
<td></td>
<td>(0.0585)</td>
<td>(0.0648)</td>
<td>(0.1496)</td>
<td>(0.0647)</td>
<td>(0.1495)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*g(t-4, t-1)</td>
<td>-0.0001</td>
<td>-0.0711</td>
<td>0.1566**</td>
<td>-0.0737</td>
<td>0.1594**</td>
</tr>
<tr>
<td></td>
<td>(0.0482)</td>
<td>(0.0790)</td>
<td>(0.0769)</td>
<td>(0.0798)</td>
<td>(0.0779)</td>
</tr>
<tr>
<td>g(t-4, t-1)</td>
<td>0.2357</td>
<td>0.4295*</td>
<td>-0.8995*</td>
<td>0.3833*</td>
<td>-0.8567</td>
</tr>
<tr>
<td></td>
<td>(0.2057)</td>
<td>(0.2231)</td>
<td>(0.5303)</td>
<td>(0.2243)</td>
<td>(0.5293)</td>
</tr>
<tr>
<td>Home Wealth (t-4)</td>
<td>0.0650***</td>
<td>0.0660***</td>
<td>0.0177</td>
<td>0.0212</td>
<td>0.1678***</td>
</tr>
<tr>
<td></td>
<td>(0.0163)</td>
<td>(0.0163)</td>
<td>(0.0280)</td>
<td>(0.0271)</td>
<td>(0.0236)</td>
</tr>
<tr>
<td>Smart &amp; Illicit</td>
<td>-0.0853</td>
<td>-0.2718**</td>
<td>0.5244**</td>
<td>-0.2717**</td>
<td>0.5296**</td>
</tr>
<tr>
<td></td>
<td>(0.1197)</td>
<td>(0.1322)</td>
<td>(0.2629)</td>
<td>(0.1322)</td>
<td>(0.2630)</td>
</tr>
<tr>
<td>Home Wealth(t-4)*ΔUnemployment</td>
<td>-1.9176</td>
<td>0.1322</td>
<td>0.2629</td>
<td>0.1322</td>
<td>0.2630</td>
</tr>
<tr>
<td></td>
<td>(1.2510)</td>
<td>(1.9307)</td>
<td>(1.7988)</td>
<td>(5.5172)</td>
<td>(12.1840)</td>
</tr>
<tr>
<td></td>
<td>(5.0715)</td>
<td>(5.1512)</td>
<td>(12.1840)</td>
<td>(5.1512)</td>
<td>(12.1840)</td>
</tr>
<tr>
<td>Observations</td>
<td>93722</td>
<td>93722</td>
<td>93755</td>
<td>93755</td>
<td>93755</td>
</tr>
</tbody>
</table>

TABLE IX
SELECTION ON WAGES AND SHOCKS TO HOME WEALTH: ADDITIONAL CONTROLS
Notes: This table reports logit (columns 1-2) and multinomial logit (columns 3-6) analyses of selection into different self-employment, unincorporated self-employment, and incorporated self-employment on pre-determined salaried wages (Wages (25-29)) and a Bartik instrument for changes in home wealth (Home Wealth(t-4)*g(t-4, t-1)). The analyses are similar to Table VII, except that the analyses in this table also control for Smart & Illicit (as well as AFQT and Illicit individually), where Smart & Illicit equals one if AFQT is 50 or above and Illicit is zero or above and Smart & Illicit equals zero otherwise. In columns (1-2), the dependent variable is a one-zero indicator variable of whether the individual is self-employed in year t. For the dependent variable in columns (3-6), we use one-zero indicator variables of employment type, where the reported categories are unincorporated and incorporated, and the unreported categories are salaried, unpaid family, and other business ownership. For the Bartik instrument, 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 and Home Wealth(t-4) is the individual's net home wealth in year t-4. As indicated, some regressions control for the change in the state unemployment rate between t-1 and t (∆Unemployment) and its interaction with Home Wealth(t-4). 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. The logit regressions also include gender-year and race-year effects. The sample is restricted to sample years 1989-2014 and excludes individuals who were self-employed in either t-2 or t-4. Table A1 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.
Figure 1: Amount Used to Establish/Acquire the Business

1. **Unincorporated**: Sole proprietorship
2. **Incorporated**: Partnership or limited liability; Limited liability corporation; Sub-chapter S corporation; General corporation
3. **Others**: Nonprofit organization; Other (specify)
Figure 2: AFQT Score by Amount Used to Establish/Acquire the Business

1. **Unincorporated**: Sole proprietorship
2. **Incorporated**: Partnership or limited liability; Limited liability corporation; Sub-chapter S corporation; General corporation
3. **Others**: Nonprofit organization; Other (specify)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human capital</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>Illicit</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>Smart &amp; Illicit</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>Rosenberg self-esteem (standardized)</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>Rotter locus of control (standardized)</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>Years of schooling</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 is 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 details.

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.

5. Starting Capital, Sources of Capital, Employees

Starting Capital
Respondent reports how much money was used to establish or acquire the business. This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Family saving
Respondent reports that personal or family savings were used to establish or acquire the business (or not). This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Loan:
business/family/friend
Respondent reports that business loans or investments from family or friends were used to establish or acquire the business (or not). This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Home equity
Respondent reports that a personal or family home equity loan was used to establish or acquire the business. (or not). This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Loan: bank/finance
Respondent reports that a loan from a bank or other financial institution was used to establish or acquire the business. (or not) This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Credit card
Respondent reports that a business or personal credit card was used to establish or acquire the business. (or not) This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Government
Respondent reports that a business loan from federal, state, or local government was used to establish or acquire the business. (or not) This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Other
Respondent reports that some other source of finance was used to establish or acquire the business. (or not) This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

None needed
Respondent reports that no money was need to establish or acquire the business. (or not) This information is obtained from the business ownership part of the NLSY79 that was given in survey years 2010 and 2012.

Employees
The number of employees includes all paid employees in the year that the person becomes full-time self-employed and excludes the self-employed business owner, which is available from 2002 onwards in the NLSY79.
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.

\( \Delta \text{Unemployment} \)  
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 in 1979. We use the cross-sectional (6,111 individuals), the supplemental (5,295 individuals), and military (1,280 individuals) samples.

(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.