Entrepreneurship, Human Capital, and the Misallocation of Talent

Alexandros Loukas †
Washington University in St. Louis

June 2023

Abstract

Is the allocation of human capital between entrepreneurs and workers a key determinant of aggregate productivity and income? How pervasive are its implications for macro-development? Analyzing cross-country population survey data reveals a new empirical fact: the mean-adjusted rate of entrepreneurship for higher educated individuals varies positively and significantly with output per worker. Focusing further on the US, I find an asymmetric U-shaped relationship between the entrepreneurship rate and educational attainment, with the left branch declining over time. Entrepreneurial outcomes such as business income and firm employment size are strongly increasing in the educational level of the firm owner/manager. Results are robust upon conditioning on a rich set of observables. I rationalize these findings within a dynamic general equilibrium framework, in which heterogeneous agents make schooling and occupational choices. Under the hypothesis that human capital can enhance the effective productivity of entrepreneurs via costly technology adoption, theory suggests that the entrepreneurship-education nexus has first-order consequences for macro-development. Calibrations of the model economy indicate potentially sizeable misallocation losses vis-à-vis the US due to inadequate complementarity between entrepreneurial talent and human capital. The ramifications are even more far reaching when considering variations in educational attainment, as a lower stock of human capital can depress both factor accumulation and aggregate TFP formation.

JEL codes: E23, E44, J24, L26, O11, O15, O49.

Keywords: Entrepreneurship; Human capital; Productivity; Complementarity; Misallocation; Macro-development.

PRELIMINARY AND INCOMPLETE: PLEASE DO NOT CIRCULATE

† Department of Economics, Washington University in St. Louis, 1 Brookings Dr, Saint Louis, MO 63130 (e-mail: alex.loukas@wustl.edu)

Acknowledgments: I am grateful to Costas Azariadis and Ping Wang for valuable comments and encouragement. I have also benefited from discussions with Daron Acemoglu, Paco Buera, Yannis Ioannides, Costas Meghir, Ben Moll, Werner Ploberger, and B. Ravikumar. The usual disclaimer applies.

I. Introduction

Entrepreneurship is widely recognized as a key driver of economic development for more than a century. At least since the pivotal works of Schumpeter (1911) and Knight (1921), a large array of research linking the intensity and quality of entrepreneurship to economic expansion has emerged. Accordingly, *who* becomes an entrepreneur is of paramount significance. Occupational choice determines and is being determined by macroeconomic outcomes, and in turn, the allocation of factors and endowments between occupations shapes total factor productivity (TFP) and output.

In an extensive body of literature, the decision to start and run a firm usually boils down to differences in entrepreneurial ability/talent; wealth and access to credit; risk aversion; or all of the above. Nevertheless, there is relatively scarce research focusing on differences stemming from more conventional human capital, i.e., related to the Schultz-Becker-Mincer view. This paper aims to first fill in this gap, and then ask a more fundamental macroeconomic question: Is the allocation of human capital between entrepreneurs and workers a key determinant of aggregate productivity and income? If so, how pervasive are its implications for macro-development?

The study has two primary objectives. As a starting point I seek to establish whether the allocation of human capital, as proxied by attained formal education by entrepreneurs and workers, is tightly linked to long-run development. For this purpose I compile an extensive dataset covering nearly 100 countries drawn from Global Entrepreneurship Monitor (GEM) adult population surveys. To address the question in simple quantitative terms, I introduce the ERGON index (Entrepreneurship Rate of Graduates Over National average), which is the normalized/mean-adjusted rate of entrepreneurship for higher educated individuals. Merging the GEM dataset with PWT 10.1 data reveals a new empirical fact: the ERGON index varies positively and significantly with output per worker. This finding is robust even after controlling for variables that are known to be firmly connected with economic advancement, such as the overall rate of entrepreneurship and measures of aggregate human capital or average years of schooling.

Focusing further on the U.S. economy, I use nationally representative samples drawn from eleven waves of the Survey of Consumer Finances (SCF) to analyze a number of prominent features related to active firm owners/managers. I document large and persistent differences in ample economic and demographic characteristics across entrepreneurs and workers in the U.S. labor force, as well as within entrepreneurs. Special attention is given to variations stemming from human capital/completed formal education. I find the relationship between the rate of entrepreneurship and educational attainment to be nonlinear and time-varying, exhibiting an asymmetric U-shape with its left branch declining over time. Among entrepreneurs, higher education is strongly positively associated – not U-shaped – with numerous measures of business outcomes such as pre-tax profits (both hourly and total), sole proprietorship income, firm employment size, net business value, and asset holdings. Conditioning on a rich set of observables, I confirm the robustness of the descriptive results through a series of reduced-form regressions under proper repeated imputation inference.

The second objective is to develop a versatile dynamic general equilibrium framework that can jointly explain the conclusions of the empirical analysis and be broadly consistent with aggregate and survey data. The ultimate goal is to shed some light on the importance of the allocation of human capital between entrepreneurs and workers for macro-development and aggregate TFP formation. To that end, I propose an elaboration of a heterogeneous agent model that features occupational choice, endogenous schooling, and incomplete markets along with the option of some financial frictions. I claim that incorporating this dimension of human capital together with my elaborations to the workhorse model would be beneficial to the relevant literature.

At the crux of most theoretical results lies the following hypothesis: there exists a disembodied technology that coalesces entrepreneurial talent and human capital to form the fundamental "effective productivity" that is used in the production process. Active firm managers face a costly intratemporal technology choice, which enables them to complement their talent with the skills and competences that education entails. Put simply, entrepreneurial human capital may serve as an additional factor of production. Competitive equilibria encompass threshold educational levels beyond which agents choose to adopt the technology; the shape and form of the entrepreneurship-education nexus is subsequently linked to aggregate economic activity.

I argue that this is not only a reasonable but also an empirically useful postulation, as the ensuing threshold levels can be inferred from population survey data using binary response regressions, or directly estimated using structural methods. Chiefly, the proposed theory is valid under the assumption of independence between entrepreneurial talent and human capital at the population level, and in fact holds for any sequence of joint distributions. If they were indeed taken to be unconditionally related in a specific way, the answer would be far from interesting.

Calibrated versions of the model economy are successful in replicating various salient features of U.S. and international data: e.g., substantial wealth and income heterogeneity with the appropriate concentration in the hands of entrepreneurs, large dispersion in business outcomes that match firm size distributions, empirically relevant cross-sectional variation in educational attainment, and reasonable firm dynamics, among others. Importantly, results are remarkably close to the findings of the empirical section: the relationship between education and entrepreneurship arises indeed as an asymmetric U-shaped form, and the theory is flexible enough to accommodate several shapes and forms. A similarly good fit applies to the relationship between human capital and entrepreneurial income, wealth, and hired capital and labor. The analysis focuses on stationary competitive equilibria, the existence and uniqueness of which is established

Mean Field Game theory.The quantitative model is solved numerically using an implicit upwind finite difference scheme following the methods described in Achdou et al. (2021).

The rest of the paper is outlined as follows. In Section II I briefly review the most relevant theoretical and empirical literature and highlight connections with the current study. Section III contains the main empirical contributions. Section IV lays out the model economy, presents a number of theoretical results which pin down the conditions that shape household occupational choices, and offers insights upon the impact of the entrepreneurship-education nexus on macro-development. Section V carries out an extensive quantitative exploration via calibration of the model economy. Section VI contains various counterfactuals . Section VII concludes the analysis and offers final comments together with a few suggestions for further work.

II. Related Literature $[INCOMPLETE]$

The seminal papers of Lucas (1978) and Rosen (1982) set up models of occupational choice, in which individuals with better managerial skills will opt to become entrepreneurs if the expected returns from starting their own firm exceed the wage they can earn as workers. Parallel to those efforts, the influential study of Kihlstrom and Laffont (1979) emphasized the relevance of risk aversion in choosing between occupations and its subsequent general equilibrium effects. Later contributions by Evans and Jovanovic (1989), Banerjee and Newman (1993), Aghion and Bolton (1997), and Piketty (1997) to name a few, paved the road to the modern literature by stressing the importance of wealth heterogeneity and financial frictions in the process of development.

In terms of quantitative theory, this study builds on a number of antecedents. The dynamic general equilibrium approach I follow is related to a series of more recent "third-wave" papers, with some noteworthy contributions being Cagetti and De Nardi (2006), Banerjee and Moll (2010), Buera, Kaboski and Shin (2011), Buera and Shin (2013), Midrigan and Xu (2014), and Moll (2014), among others.[1]

In light of the abundance of data during the past decades, one should be tempted to ask *who becomes an entrepreneur and what makes a successful one?*. Some empirical answers have been provided by Hamilton (2000), Cagetti and De Nardi (2006), and Poschke (2013) for the case of the US, and by Blanchflower (2000) and Blanchflower (2004) for the case of OECD countries.

III. Empirical Analysis

Not all entrepreneurs are created equal. Heterogeneity matters, and it matters in various ways. This section is an empirical study that throws light on the sources of heterogeneity across entrepreneurs and wage earners, as well as within entrepreneurs. The first

[1] As this literature is too extensive to cover, please see the excellent surveys of Quadrini (2009) and Buera, Kaboski and Shin (2015).

contribution concerns the documentation of salients features of active entrepreneurs in the U.S. over the last three decades using a wide-ranging set of demographic and personal characteristics, with special focus on years of formal education. The second contribution pertains to the analysis of selection into entrepreneurship along with systematic variations in business outcomes stemming from differences in educational attainment, which is interpreted as a primary measure of human capital.

III.1. U.S. Dataset Description and Methodology

I use nationally representative data collected by the Survey of Consumer Finances (SCF) over the period $1989-2019$. The SCF is a detailed triennial cross-sectional survey of U.S. households sponsored by the Board of Governors of the Federal Reserve System in cooperation with the Statistics of Income Division of the IRS. I do not consider the two datasets prior to 1989 (1983 and 1986) as those surveys are known to be of lower overall quality, missing data were singly and not multiply imputed, and chiefly, interviewers did not ask a number of questions regarding the households' business activities.

Samples in the SCF are drawn based on a dual-frame sampling scheme. The area-probability sample is selected in three stages in order to provide robust national coverage of a broad range of behaviors and characteristics, and the list sample is selected using information from administrative data in order to disproportionately sample more wealthy households. One important advantage is having many nationally-representative samples that adequately capture demographic and occupational characteristics, as well as the full distribution of wealth and income. The SCF design also benefits the purpose of this study since the over-sampling of the wealthy results in more observations of households associated with entrepreneurship.

There is a number of methodological aspects that need to be taken into account, especially those related to multiple imputation and repeated imputation inference (RII). The theory on proper inference in this setting is well-understood; it was developed by Donald B. Rubin in the 1970's and exemplified in his famous book Rubin (1987) (see also the survey of Little (1992), as well as Van Buuren (2018) for a good example of more recent work in the field). In Appendix C I summarize what some of these issues are and explain how the current study addresses them.

III.2. Who Are the Entrepreneurs and How They Are Identified

A nontrivial question is how to identify entrepreneurs in survey data. The approach I follow is inspired by Cagetti and De Nardi (2006) and De Nardi, Doctor and Krane (2007), but instead of deciding solely upon the answers of the arbitrarily defined "household head", I take into account both the respondent (R) and the spouse/partner (S/P).[2]

I classify as entrepreneurs the households in which either Rt or S/P meets the following four criteria: **(i)** engages in some form of self-employment as their primary occupation,[3] **(ii)** owns or shares ownership in at least one privately-held business, **(iii)** has an active management role in at least one business, and **(iv)** the net value of actively managed businesses in greater than zero. Throughout the paper I shall refer to such households as *self-employed business owners (SEBO)*, or simply as *entrepreneurs*. Likewise, I shall refer to the rest of the labor force as *non-SEBO*, or simply as *workers/wage-earners*.

The above classification is fairly noncontroversial since the mapping of related models to data calls for entrepreneurs to have an investment stake in their business and to be working as managers with some span of control over hired capital and labor. The requirement of self-employment eliminates people who are predominantly employed by someone else and only help in a business as mere pastime. The requirement of owning (part of) the business helps precluding counting as entrepreneurs people who are temporarily self-employed or switching between jobs. The requirement of a managerial role in at least one firm phases out potentially wealthy individuals who may participate in a business only as passive investors. Also note that this definition does not discriminate against the legal status of business, i.e., sole proprietorship, unincorporated partnership, subchapter S or C, or an LLC/LLP company. More details can be found in Appendix C.

III.3. Descriptive Analysis

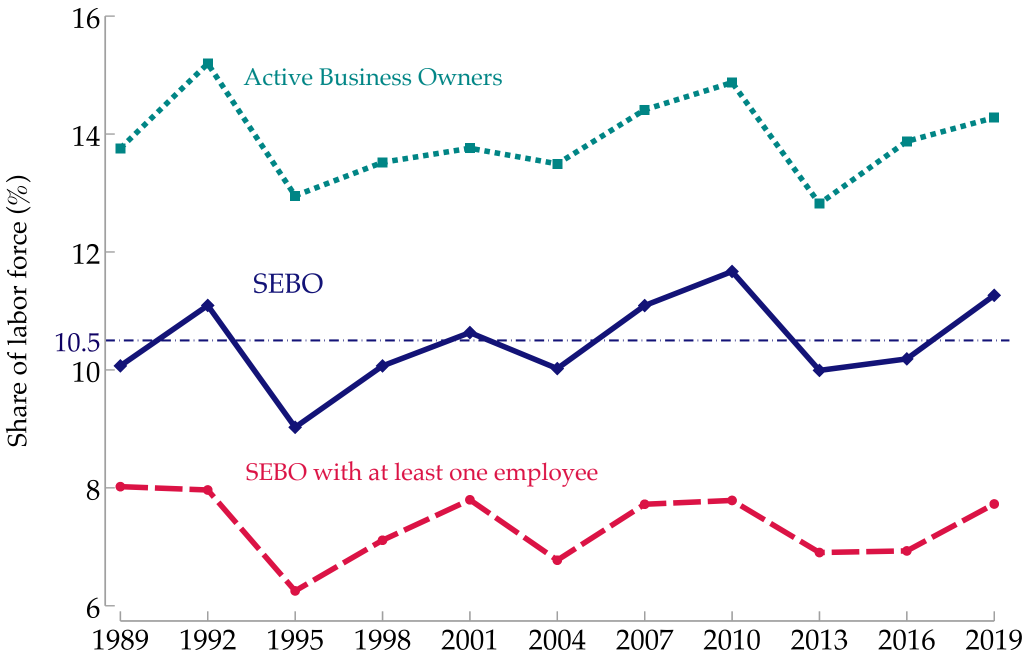
Figure 1 shows the rate of entrepreneurship in the U.S. labor force according to three different definitions, with SEBO being the adopted measure throughout this study. The fraction of entrepreneurs has remained relatively stable over the thirty years, hovering around 10.5 percent of the labor force. This average value is markedly close to the arithmetic mean of 10.4 percent for the U.S. entrepreneurship rate over $1990 - 2009$, as inferred from Current Population Survey (CPS) data reported by Hipple (2010).[4] The apparent consistency with U.S. Census/BLS data lends further credence to the procedures

[2] Although most empirical analyses using the SCF have considered the personal characteristics of the "head of the PEU", it is important to realize that the respondent and the head in couple households are not necessarily interchangeable. R is identified by the SCF staff in the initial screening interview as the more financially knowledgeable person. For mixed-sex couple households, the SCF always assigns the title to the male partner, and for same-sex couple households the older person in the partnership is assigned that title. There is no reason why to simply assume that the head's demographic and personal characteristics (e.g., sex, race, age, experience) correspond to those of the SEBO in the household.

[3] To be more precise, the SCF provides the following answers relating to self-employment status: i) Self-employed; other closely held business owned by PEU; ii) Partnership; law firm; medical/dental partnership; other non-publicly traded business in which R/S/P has an interest; iii) Consultant / Contractor.

[4] Specifically, this number corresponds to the sum of the incorporated plus unincorporated self-employed divided by total employment in all non-agricultural industries; see Table 1 and Table 2 in Hipple (2010). The results are identical to the first decimal point whether calculating the average of ratios or the ratio of averages.

Figure 1: Fraction of Entrepreneurs in the U.S. Labor Force



Sources: Survey of Consumer Finances and author's calculations.

I employ and to the argument for truly nationally representative empirics.

Entrepreneurs may form a relatively small group compared to the working population, but their economic significance is disproportionately large. Two of the most conspicuous examples are depicted in Figure 2 and Figure 3. Despite being on average only 10.5 percent of the labor force, they hold 44 percent of total net worth. The systematic concentration of wealth in the hands of SEBO becomes even more apparent when we look at higher percentiles of the distribution. About 4 in 10 working adults with net worth in the top decile are entrepreneurs; around 1 in 2 in the top 5 percent; and this number climbs to almost two thirds when considering the top 1 percent of wealth holders. Furthermore, entrepreneurs receive on average between one fifth and one fourth of total U.S. labor force income. They are also substantially overrepresented in the top percentiles of the income distribution, for example by a factor of 3 and 5 when looking at the top decile and top 1 percent, respectively. However, it is worth stating that these numbers pertaining to entrepreneurial income should be taken more as lower bounds rather than definitive estimates, as it has been known in the literature (e.g. Hamilton (2000)) that legally disclosed net business profit may be deceptively low because of various tax incentives to underreport it to the IRS.

Figure 2: Share of U.S. Labor Force Wealth Held by Entrepreneurs (Left Panel); Fraction of Entrepreneurs in Selected Wealth Percentiles (Right Panel)

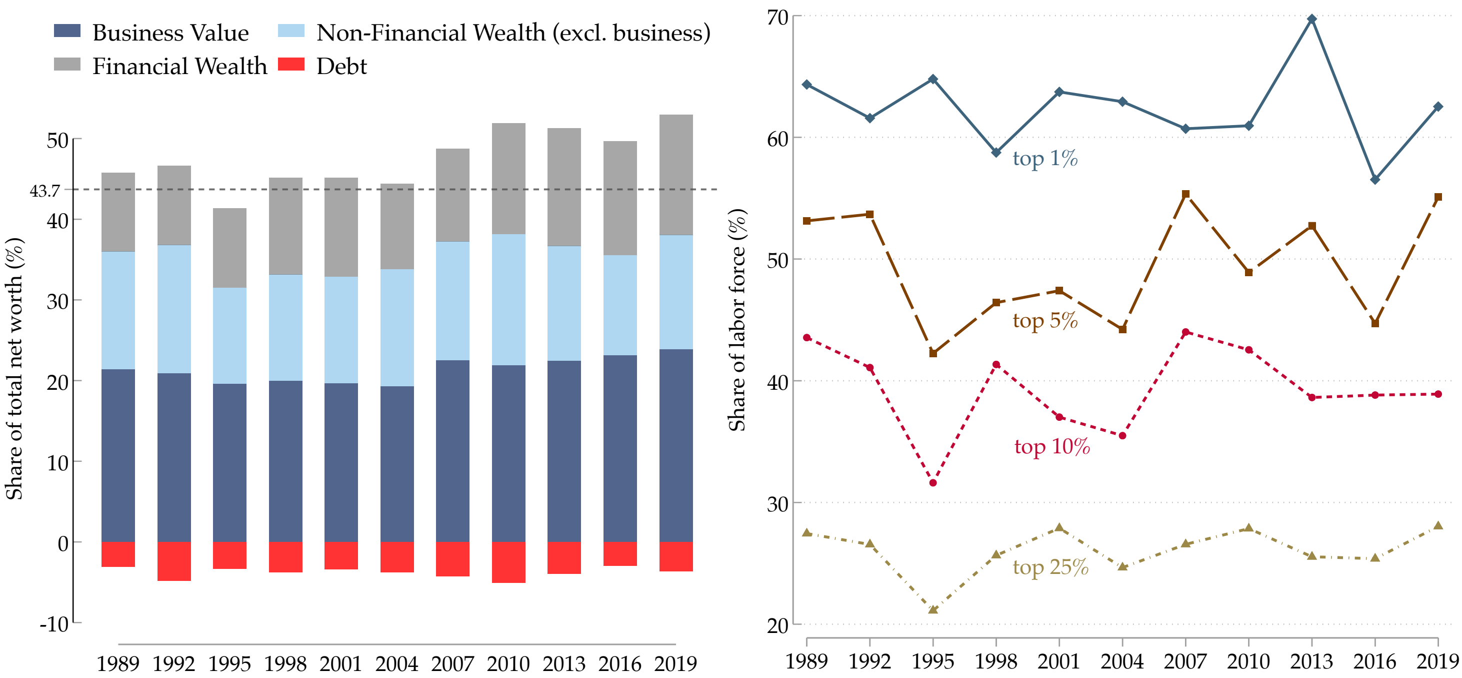
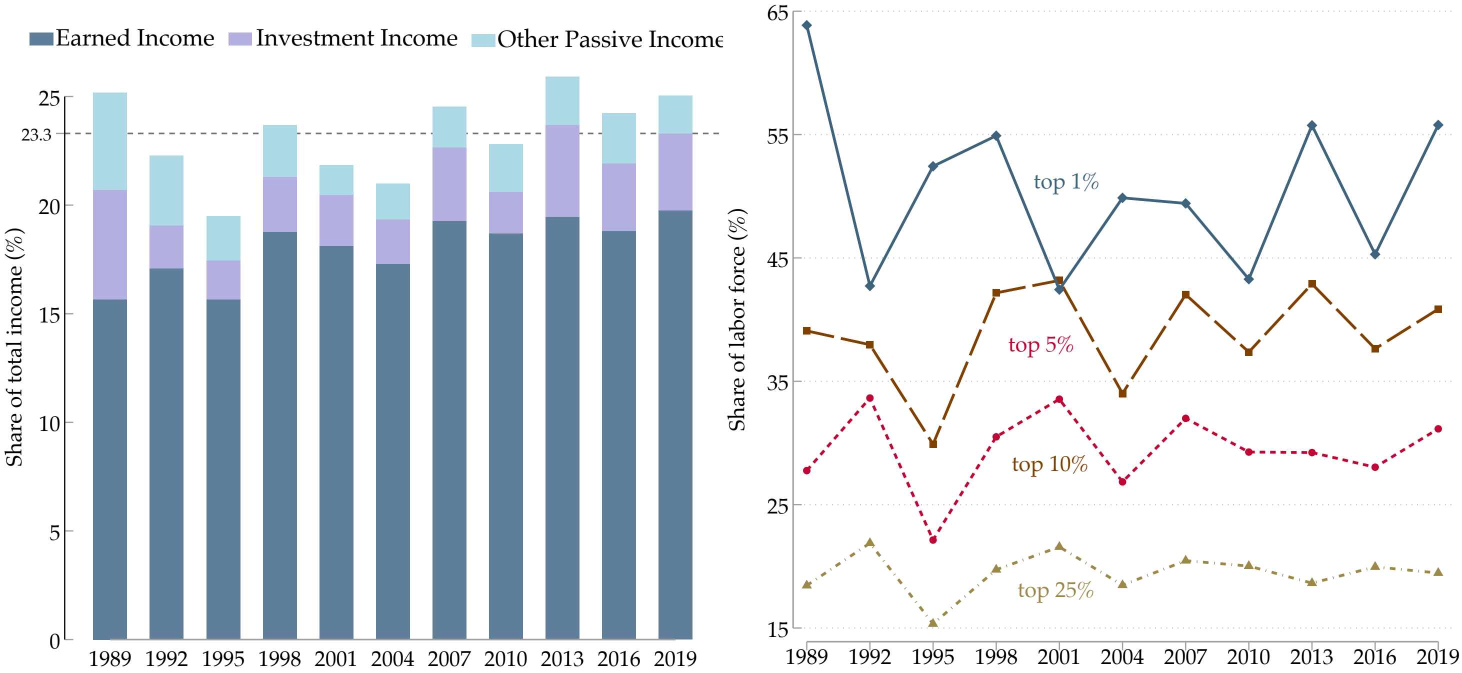


Figure 3: Share of U.S. Labor Force Income Held by Entrepreneurs (Left Panel) ; Fraction of Entrepreneurs in Selected Income Percentiles (Right Panel)



Sources (both figures): Survey of Consumer Finances and author's calculations.

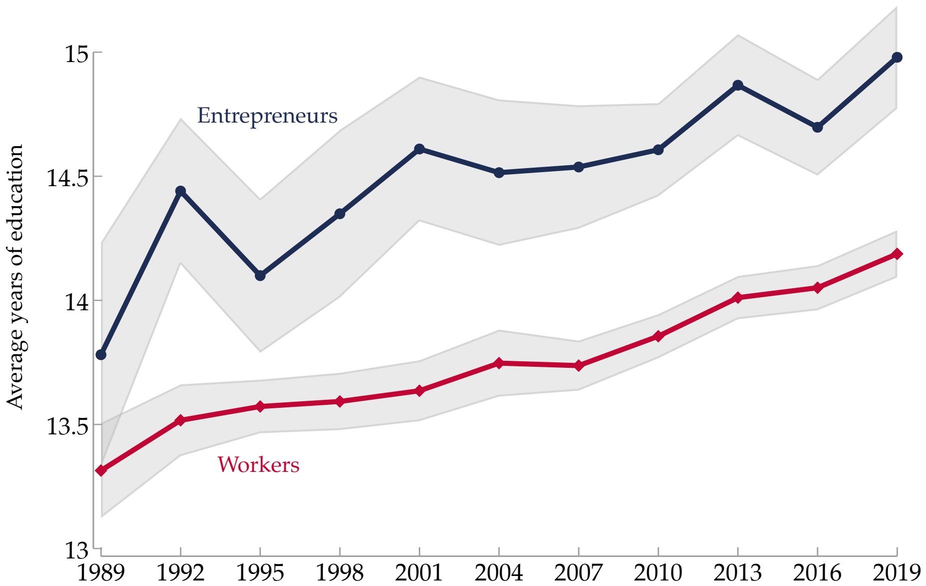
Table 1 documents ample differences in various personal and demographic characteristics across entrepreneurs and workers in the U.S. labor force. In a nutshell: entrepreneurs are more educated than workers; they have more years of potential experience in the labor market – they are also older on average; they work more and for longer on a yearly basis; they score better when it comes to self-assessing their general health condition; and they report higher willingness to take financial risk. There seem to be no systematic differences in their rates of retirement or physical/mental disability. In addition, they are predominantly married (or in a civil union), male, and White/Caucasian, although this has been steadily changing over the 21st century in line with demographic shifts. For example, Black/African-American and Hispanic/Latino populations that have been underrepresented in the entrepreneurial group are catching up in the second part of the subsample, whereas the Other Race category, which consists of mostly Asian people, has always been equally represented in both categories. Undoubtedly many of the above disparities between occupational groups deserve a deeper look, and many of them have been already studied to some extent in the relevant literature, but the dimension of education has been largely overlooked. This is the topic I shall now turn to.

The evolution of average educational attainment for entrepreneurs and workers is shown in Figure 4. The mean differences are extensive and somewhat widening over time, with the plotted 95% RII confidence intervals suggesting very high statistical significance. A discrepancy of $0.8 - 0.9$ years of schooling on average should be considered abundant, not only within a country's labor force, but also across different countries.

To gain some perspective on what this average difference translates to in international data, I consult the latest version of the Barro-Lee educational attainment dataset for population aged 25-64. Over the period $1995 - 2015$, if we look at average years of schooling in some of the world's richest advanced economies such as Ireland, Denmark, Sweden, Germany, Australia, and Japan, and compare them to some less advanced and/or upper-middle income countries such as Hungary, Poland, Lithuania, Latvia, Kazakhstan, and Ukraine, the differences are less or equal to 0.9 years of schooling. A difference of 0.5 years or less can be obtained if we look at some of the relatively successful African countries such as Tunisia, Egypt, and Namibia, and compare them to much poorer countries like the Republic of the Congo, Syria, and Cameroon.[5]

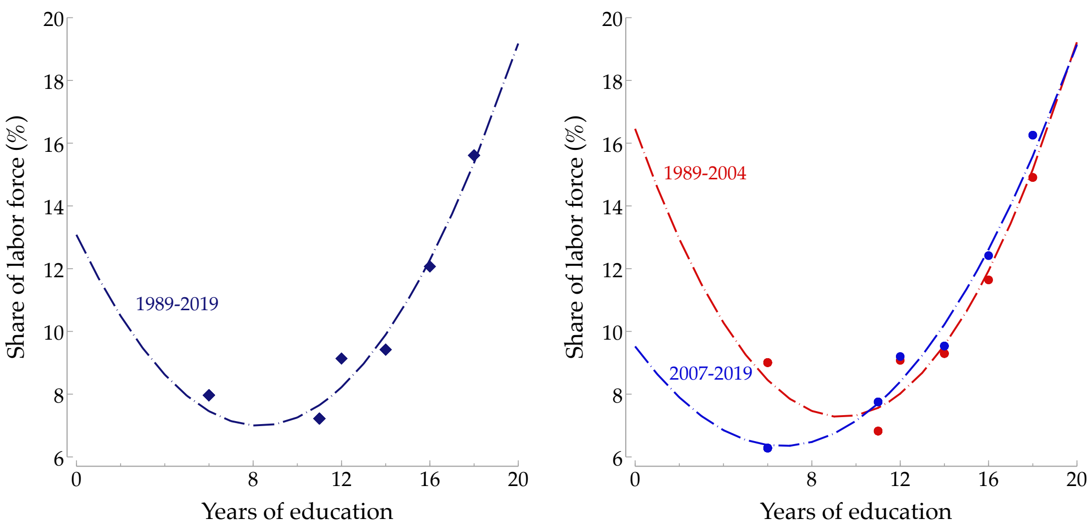
[5] For completeness, these are the underlying statistics (time averages for $1995 - 2015$). Denmark: 12.1; Sweden: 12.1; Japan: 12.1; Germany: 12.0; Ireland: 12.0; Australia: 12.0; compared to, Hungary: 11.9; Ukraine: 11.7; Poland: 11.4; Lithuania: 11.3; Latvia: 11.2; Kazakhstan: 11.2. We also have, Tunisia: 6.3; Namibia: 6.3; Egypt: 6.2; compared to, the Republic of Congo: 5.9; Syria: 5.9; Cameroon: 5.8.

Figure 4: Average Years of Education: U.S. Entrepreneurs vs Workers



Sources: Survey of Consumer Finances and author's calculations.
Notes: Connected markers refer to weighted arithmetic means by occupational group. Sample means and standard errors are calculated under RII using all 5 SCF implicates for every observation and all 999 SCF bootstrap replicate draws and weights. Shaded areas represent 95% confidence intervals for the mean.

Figure 5: U.S. Entrepreneurship Rate by Educational Attainment



Sources: Survey of Consumer Finances and author's calculations.
Notes: The graph depicts the fraction of entrepreneurs in six different educational attainment categories, along with a fitted quadratic regression line using RII; full pooled sample (left panel) and two consecutive subsamples (right panel).

Table 1: Descriptive Statistics: Entrepreneurs vs Workers

Variable	Pooled 1989–2004 Sample			Pooled 2007–2019 Sample		
	non-SEBO (N_1=13,932)	SEBO (N_2=5,187)	p-value of difference	non-SEBO (N_1=16,601)	SEBO (N_2=5,502)	p-value of difference
Education	13.6	14.4	$<.01$	14.0	14.9	$<.01$
Potential experience	22.3	26.5	$<.01$	24.2	30.7	$<.01$
Worker experience	21.2	15.7	$<.01$	23.2	17.4	$<.01$
Self-employed experience	1.1	10.8	$<.01$	1.0	13.3	$<.01$
Annual labor supply	$2,142$	$2,418$	$<.01$	$2,109$	$2,230$	$<.01$
Health $(1-4)$	3.19	3.31	$<.01$	3.08	3.20	$<.01$
Risk willingness $(1-4)$	1.87	2.14	$<.01$	1.88	2.17	$<.01$
Employment variety	3.15	2.98	$<.05$	3.34	3.08	$<.05$
Unemployed partner $(0/1)$	0.25	0.22	$<.01$	0.26	0.23	$<.01$
Retired $(0/1)$	0.007	0.010	$<.10$	0.02	0.03	$<.10$
Disabled $(0/1)$	0.002	0.003	.65	0.010	0.008	.61
Male	0.68	0.77	$<.01$	0.65	0.73	$<.01$
Married	0.65	0.82	$<.01$	0.63	0.82	$<.01$
White/Caucasian	0.74	0.89	$<.01$	0.64	0.79	$<.01$
Black/African-American	0.12	0.03	$<.01$	0.14	0.05	$<.01$
Hispanic/Latino	0.09	0.03	$<.01$	0.12	0.06	$<.01$
Other non-white	0.05	0.05	.96	0.10	0.10	.79

Notes: All statistics refer to weighted arithmetic means. Sample means and standard errors are calculated under RII using all 5 SCF implicates for every observation and all 999 SCF bootstrap replicate draws and weights. The equality of means/proportions is assessed with a weighted, two-sided, unequal-variance hypothesis test.

Table 2: Within-Entrepreneur Heterogeneity by Educational Attainment

Variable	Years of Education (pooled 1989–2019 Sample)					
	[0,8]	(8,12)	12	(12, 16)	16	(16, 21]
Total business income (> 0)	53,302	64,625	82,502	93,122	188,941	260,210
Total business income	38,831	45,332	54,886	64,102	129,745	195,047
Sole proprietorship income	26,014	33,849	38,948	42,098	59,181	100,865
Total business income per hour (> 0)	23.4	35.1	54.9	64.5	114.5	229.0
Total business income per hour	17.1	24.5	37.0	42.0	78.6	170.5
Sole proprietorship income per hour	12.5	19.9	28.1	33.3	34.2	60.2
Firm employment size (> 0)	5.4	6.3	8.2	16.8	49.1	80.9
Firm employment size	3.4	4.1	5.7	10.9	35.5	62.1
Firm net value per employee	85,725	90,667	137,249	140,997	209,560	233,104
Main business ownership share	0.92	0.91	0.91	0.89	0.82	0.77
Potential experience	35.1	33.4	30.7	27.8	26.9	27.5
Self-employed experience	14.6	13.6	12.7	11.0	11.1	12.8
Prior worker experience	20.5	19.8	18.1	16.8	15.7	14.8
Risk willingness ($1-4$)	1.59	1.66	1.94	2.18	2.36	2.37
Health ($1-4$)	3.32	3.28	3.35	3.37	3.50	3.52
Ever received inheritance ($0/1$)	0.18	0.17	0.26	0.30	0.34	0.36
Expect to receive inheritance ($0/1$)	0.10	0.12	0.16	0.22	0.26	0.25
Male	0.92	0.82	0.76	0.68	0.72	0.77
Married	0.89	0.77	0.85	0.79	0.83	0.82
White/Caucasian	0.66	0.75	0.84	0.83	0.85	0.86

Notes: All statistics refer to weighted arithmetic means, calculated under RII using tall 5 SCF implicates for every observation and all 999 SCF bootstrap replicate draws and weights.

III.4. *Reduced-form Analysis* [$INCOMPLETE$]

The descriptive results show dramatic differences along the dimension of educational attainment, whether we consider occupational choice between entrepreneurship and wage employment, or business outcomes within entrepreneurs. It is nonetheless possible that these differences stem from a variety of sources aside from education. It could be the case, for example, that highly educated individuals work longer hours and have more job market experience, or that more educated entrepreneurs are on average healthier and more risk-loving. In such cases the apparent impact of education could be fairly biased upward, so one would want to control for as many confounding variables as possible.

In order to further sort out the effect of formal education on the rate of entrepreneurship and on entrepreneurial outcomes, this section presents reduced-form results (probit and OLS) for pertinent outcome variables. I consider regressions of the following form:[6]

$$y_i = \beta_0 + \beta_1 \cdot educ_i + \beta_2 \cdot educ_i^2 + X_i\boldsymbol{\gamma} + \Gamma_s + \Gamma_o + \Gamma_s \times \Gamma_o + \Gamma_t + \varepsilon_i$$

where X_i is a rich set of controls such as potential experience, annual hours worked, self-reported health condition and willingness to take risk, past and/or future inheritances, marital status, sex, and race, among others. I also control for sector fixed effects (Γ_s), occupation fixed effects (Γ_o), their interactions ($\Gamma_s \times \Gamma_o$), and year fixed effects (Γ_t).

The coefficients β_1 and β_2 are of primary interest. For the probit regressions, if the estimated coefficient $\widehat{\beta}_1$ is positive (negative) and strongly significant under conservative standard errors, it can be interpreted as evidence that higher education has a positive (negative) impact on the selection into entrepreneurship. The more interesting case is when both hypotheses $\widehat{\beta}_1 = 0$ and $\widehat{\beta}_2 = 0$ are strongly rejected so that a non-linear relationship emerges, even after controlling for a wide range of factors. The exact combination of signs and magnitudes of $(\widehat{\beta}_1, \widehat{\beta}_2)$ will determine the vertex $(-\widehat{\beta}_1/2\widehat{\beta}_2)$ and consequently the shape and asymmetry of the parabola within the support of $educ$.

The evidence presented can be also useful in addressing a famous myth surrounding business success. Reflecting on the link between education and entrepreneurial achievements, one may be (unconditionally) tempted to conclude that the two are unrelated or even inversely related. Famous examples of extraordinary entrepreneurs who happen to be college dropouts come to mind, such as Paul Allen, Michael Dell, Larry Ellison, Bill Gates, Steve Jobs, Steve Wozniak, and Mark Zuckerberg to name a few. One could go even further and consider famous high school dropouts such as Richard Branson, Amancio Ortega, and Francois Pinault. But such examples are so memorable exactly because they are so rare. All in all, very successful entrepreneurs without higher formal educa-

[6] Regressions are based on the entire pooled sample and are weighted by the appropriate SCF sampling weights; estimated parameters are based on repeated imputation inference using all 5 SCF implicates for every observation; bootstrapped standard errors are calculated using 999 SCF bootstrap replicate draws and their respective replicate weights. As before, please refer to Appendix C for more details

tion are the exception rather than the rule. They form a particular subset made by tail events, with a relatively small measure compared to the set of entrepreneurs, and this is perhaps another reason why their cases are so exceptional.

IV. The Model Economy

IV.1. Setup, Endowments, Preferences

The model describes an infinite-horizon economy set in continuous time, $t \in [0, \infty)$, under the usual conditions: a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ with information given by the typical filtration induced by a Wiener process, $W_t \in \mathbb{R}^3$. The sample space is the compact set $\Omega = (\mathcal{A} \times \mathcal{Z} \times \mathcal{H}) \subset \mathbb{R}^3_+$, generated by the joint support of the three state variables described in the next paragraphs.[7] Admissible controls are chosen from the space $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ of square-integrable $\mathcal{F}_t$-adapted processes.

There is a measure-one continuum of infinitely lived agents who are *ex-ante* heterogeneous in their initial level of *wealth*, $a \in \mathcal{A}$, *entrepreneurial productivity*, $z \in \mathcal{Z}$, and *human capital*, $h \in \mathcal{H}$. We can interpret z as an inalienable characteristic of entrepreneurial talent with a wide perspective: the efficiency in combining factors of production, the quality of business ideas and management, the ability to market the consumption good successfully, or the capacity to withstand demand shocks. Likewise, we can think of human capital h as an all-encompassing measure of notions such as the benefits of formal education, fluid intelligence, and determination.

At each point in time, individuals with wealth (a) observe their human capital (h) and entrepreneurial productivity (z), and decide whether to work for labor income contingent on their skills (wh) or to become entrepreneurs and earn profits $(\pi(h, a, z|w, r))$. All agents supply their labor inelastically to the market.

The endogenous state variable of each agent is her level of wealth, which is optimally determined by forward-looking saving behavior. The two exogenous states, z_t and h_t, evolve stochastically according to independent Itô processes of the (differential) form,

$$dz_t = \widetilde{\mu}_z(z,t)\, dt + \widetilde{\sigma}_z(z,t)\, dW_t^z \tag{1}$$

$$dh_t = \widetilde{\mu}_h(h,t)\, dt + \widetilde{\sigma}_h(h,t)\, dW_t^h \tag{2}$$

where all drift and diffusion coefficients are globally Lipschitz-continuous, hence all transition functions satisfy the Feller property. I only consider processes that admit non-singular invariant measures, so that stationary and ergodic distributions do exist.[8]

[7] There is no theoretical reason *per se* to restrict the analysis of the state space to be compact; this is motivated purely by practical reasons. Since I will ultimately solve the model quantitatively and will get to choose closed and bounded grid intervals, the sets will be *de facto* compact.

[8] For the even more general case of Feller processes, a known necessary and sufficient condition for

It is also assumed that due to technological reasons the realizations of z_t and h_t lie within the compact sets $\mathcal{Z}$ and $\mathcal{H}$, respectively, as every sample path is reflected both above and below (reflecting barriers). This natural assumption gives rise to two Neumann boundary conditions for each process; see 18 and 19 below. To emphasize the relevance of the model without presupposing any dependence between entrepreneurial talent and human capital at the population level, I take the two Wiener processes to be mutually independent, i.e., $W_t^z \perp W_t^h, \forall t \in [0, \infty)$.

Agents have standard preferences over utility flows from consumption, c_t, discounted at the common rate $\rho > 0$. I assume CRRA preferences so that the instantaneous utility of consuming the final good is $u(c_t) = \frac{c^{1-\gamma}}{1-\gamma}$. Under a physical probability measure $\mathbb{P}$, agents seek to maximize the expected current value of their objective function over the set of all admissible controls $\mathcal{C}(h, a, z)$,

$$V(h, a, z, t) = \max_{c_t \in \mathcal{C}} \mathbb{E}_t \left[\int_t^{\infty} e^{-\rho(s-t)} u(c_s) \, ds \, |\mathcal{F}_t \right] \tag{3}$$

In the absence of aggregate shocks, the state vector (a_t, z_t, h_t) completely characterizes each individual at any point in time, and the state of the economy is given by the (endogenous) joint distribution $G_t(h, a, z)$.

IV.2. Financial and Labor Markets

Financial markets are fully liquid but incomplete. Agents have access solely to non-state-contingent deposits that pay interest rate r_t. Credit transactions are settled within the same period and all individuals face a common *borrowing constraint*. The state constraint for wealth becomes $a_t \geq 0, \ \forall t \in [0, \infty)$, giving rise to a state constraint boundary condition (SCBC). In conjunction with idiosyncratic productivity shocks, the borrowing restriction accentuates the Bewley component of the model: precautionary savings lead to a higher aggregate supply of capital, thus to a lower risk-free rate.

Physical capital is the only productive asset in the economy. Entrepreneurs and households trade capital via a large number of competitive financial intermediaries that receive deposits from savers and create capital loans for firms. The equilibrium interest rate is determined endogenously through the forces of aggregate supply and demand. The zero-profit condition for intermediaries implies that the rental rate equals the user cost of capital: $R_t = r_t + \delta$.

Active producers can finance their capital expenditures (k_t) using either internal funds, or by seeking external funds from financial intermediaries, with the latter being subject to frictions. Due to limited contract enforceability, they all face a common

the existence of an invariant probability measure μ_x on $\mathbb{R}$ is $\int_{\mathbb{R}} \mathcal{A}_x f(x) \mu_x(dx) = 0$, for all càdlàg $f(.)$ in the domain of its infinitesimal generator $\mathcal{A}_x$.

collateral constraint restricting their maximum debt position (d_t) to a fraction of the entrepreneur's hired capital.[9] Specifically, $d_t \leq \vartheta k_t$, and using the definition $d_t = \max\{k_t - a_t, 0\}$ the constraint becomes,

$$k_t \leq \frac{1}{1-\vartheta}\, a_t \tag{4}$$

The parameter $\vartheta \in [0, 1]$ offers a parsimonious way to capture the degree of financial frictions in the economy, where $\vartheta = 0$ corresponds to financial autarky, and $\vartheta \to 1$ results in a frictionless capital market.

The labor market is competitive; no firm has wage-setting power as the mass of each entrepreneur is strategically negligible. Workers supply labor to firms inelastically, denominated in units of human capital, with the total units of marketable human capital resolved in general equilibrium. Under perfectly substitutability of labor inputs in a competitive market, employees receive labor income equal to $w_t h_t$. The equilibrium effective wage rate, w_t, is determined endogenously through the aggregate supply of human capital by workers and the aggregate demand for labor by entrepreneurs.

This setup captures the common assumption of efficiency units of labor. When entrepreneurs hire labor/units of human capital, the actual composition of hired workers does not matter. That is, there is no assortative matching.

IV.3. Entrepreneurs and Technology

Entrepreneurs behave competitively in product and factor markets. Revenue is generated according to an individual-specific technology that turns capital and labor into the homogeneous consumption good, which is taken as the numéraire. They hire k units of physical capital and ℓ units of human capital to maximize per-period profits.

Given a wage rate per unit of human capital (w) and a rental rate of capital (r), the entrepreneur's indirect profit function is:

$$\pi(\tilde{h}, a, z | w, r) = \max_{k,\ell \geq 0} \; y(h_e, z, k, \ell) - w\ell - (r+\delta)k \tag{5}$$

$$\text{subject to} \quad k \leq \frac{1}{1-\vartheta}\, a$$

[9] This form of static collateral constraint has been quite standard in the macro and entrepreneurship literature for more than three decades; e.g., Evans and Jovanovic (1989), Kitao (2008), Banerjee and Moll (2010), Buera and Shin (2013), and Midrigan and Xu (2014), among others. It can be derived in the context of a simple limited enforcement problem with limited liability. Upon entering a contract, entrepreneurs can renege on their obligations and embezzle their full debt position; in retaliation, lenders can seize a fraction ϑ of the firm's hired capital. In a zero-profit equilibrium, intermediaries will lend only up to the amount they can recover, meaning that $d_t \leq \vartheta k_t$. In principle, as Moll (2014) has showed, the collateral constraint can potentially take much more general forms without necessarily affecting the core results, as long as the constraint is linear in wealth.

To explore the quantitative implications of the model in detail, I expand upon Lucas (1978) and consider the production technology,[10]

$$y(z, h_e, k, \ell) = \zeta(z, h_e)\left(k^\alpha \ell^{1-\alpha}\right)^{1-\nu} \tag{6}$$

where $h_e := h_i \cdot \mathbb{1}_{entrepreneur}$ denotes the level of human capital of entrepreneur i; ν is the span-of-control parameter. As a result there are diminishing returns to scale along the expansion of a firm's activity, since it becomes increasingly difficult for the manager to exert control over the production plan and thus marginal products diminish.

The nature of the function $\zeta(z, h_e)$ is consequential. I hypothesize the existence of a *disembodied technology* that coalesces an entrepreneur's talent (z) and human capital (h_e) to form the fundamental "effective entrepreneurial productivity." The output of this process is what entrepreneurs/firms ultimately use in the production process, and it involves positive complementarity between z and h_e.[11] It need not be constant across economies or over time.

Active entrepreneurs face an intratemporal *technology adoption choice* in the beginning of every period. Reaping the underlying benefits is not guaranteed, as the complexity conjoined with adopting the technology incurs some costs. To keep things uncomplicated while retaining interesting results, I propose the following specification:

$$\zeta(z, h_e) := \begin{cases} z & \text{at no cost} \\ zh^\omega & \text{at cost } \kappa \cdot y(),\ \kappa \in (0,1) \end{cases} \tag{7}$$

where the parameter ω is of particular interest as it governs the complementary between entrepreneurial talent and human capital. The cost parameter κ implies that the proportionally constant tradeoff is denominated in (common) units of output. Moreover, the function $\zeta(z, h_e)$ is uniformly continuous on any convex domain of $\mathbb{R}^2$, and for $\omega \lesseqqgtr 1$ it exhibits diminishing/constant/marginal returns with respect to h_e upon adoption.

LEMMA 1 (TECHNOLOGY ADOPTION): Given any joint distribution $G(z, h)$:

– $\exists\, \hbar(\kappa, \omega) = \left(\frac{1}{1-\kappa}\right)^{1/\omega}$ such that all ex post capital unconstrained entrepreneurs choose to adopt the technology if and only if $h \geq \hbar$. The same applies to ex ante-ex post capital constrained entrepreneurs.

– $\exists\, \hbar^C(a, \kappa, \omega) > \hbar$ such that all ex post capital constrained entrepreneurs choose to adopt the technology if and only if $h \geq \hbar^C$.

[10] In this class of models, the specific functional form of the individual production functions presented and most of the results that will be presented, including aggregation, are easily generalizable. The key assumptions are that all firm production functions are homogeneous of the same degree in all inputs and that productivity shocks are multiplicative.

[11] In slightly more technical terms, the theorized function is assumed to be *log-supermodular.*

The ensuing technology choice has an intuitive interpretation: for low levels of human capital individual choose to fully rely on their entrepreneurial ability, whereas more educated individuals will choose to complement their talent with the skills and competences that human capital entails. *Ceteris paribus*, "more productive entrepreneur = more talented or more educated." The potential impact of financial frictions is illustrated in the second part of Lemma 1; individuals who will be capital constrained upon adoption due to insufficient assets face a higher human capital cutoff.

LEMMA 2: *Net* firm output can be expressed as:

$$\widehat{y} = z\tilde{h}\left(\widehat{k}^{\alpha}\,\widehat{\ell}^{1-\alpha}\right)^{1-\nu} \tag{8}$$

$$\tilde{h} := \underset{h\leq\hbar}{\mathbb{1}} + \underset{h>\hbar}{\mathbb{1}}\left(\frac{h_e}{\hbar}\right)^{\omega} \tag{9}$$

Optimal factor demands depend on individual state vectors (h_e, a, z) and are given by:

$$\widehat{\ell}(h_e, a, z) = (z\widetilde{h})^{\frac{1}{\nu}}\left(\frac{\alpha(1-\nu)}{\widehat{r}_t+\delta}\right)^{\frac{\alpha(1-\nu)}{\nu}}\left(\frac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{1-\alpha(1-\nu)}{\nu}} \tag{10}$$

$$\widehat{k}(h_e, a, z) = (z\widetilde{h})^{\frac{1}{\nu}}\left(\frac{\alpha(1-\nu)}{\widehat{r}_t+\delta}\right)^{\frac{1-(1-\alpha)(1-\nu)}{\nu}}\left(\frac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{(1-\alpha)(1-\nu)}{\nu}} \tag{11}$$

The entrepreneur-specific shadow interest rate is defined as $\widehat{r}_t(h_e, a, z) = r_t + \lambda(h_e, a, z)$ *where* $\lambda(.) \geq 0$ *is the Lagrange multiplier on the collateral constraint, and corresponds to:*

$$\widehat{r}_t(h_e, a, z) = \begin{cases} r_t \ \text{ if } \ a > (1-\vartheta)\,(z\widetilde{h})^{\frac{1}{\nu}}\left(\dfrac{\alpha(1-\nu)}{r_t+\delta}\right)^{\frac{1-(1-\alpha)(1-\nu)}{\nu}}\left(\dfrac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{(1-\alpha)(1-\nu)}{\nu}} \\ \left(\dfrac{(z\widetilde{h})^{\frac{1}{\nu}}}{\frac{1}{1-\vartheta}a}\right)^{\frac{\nu}{1-(1-\alpha)(1-\nu)}}\left(\dfrac{(1-\alpha)(1-\nu)}{w_t}\right)^{\frac{(1-\alpha)(1-\nu)}{1-(1-\alpha)(1-\nu)}}\alpha(1-\nu)-\delta \quad \text{otherwise} \end{cases} \tag{12}$$

In the presence of a certain wealth heterogeneity in the economy, Lemma 2 establishes that collateral constraints will alter the production scale and profits of some entrepreneurs. This shows up in the *shadow cost of funds*[12] each capital-constrained producer is facing, i.e., through the size of the multiplier $\lambda(h_e, a, z)$. As the demand for capital and labor is strictly increasing in z and $\widetilde{h}$, agents on the higher end of both distributions are being disproportionately affected. The dispersion of marginal products thus creates a persistent case of capital and talent misallocation on the intensive margin.

[12] This is the terminology that Midrigan and Xu (2014) use and their work contains similar results.

All production takes place intraperiod, with earnings distributed to entrepreneurs accordingly. As individuals observe their state vector (h, a, z) at the start of each period, there can be no default. At the end of the period agents make their consumption/saving decisions. To keep things simple I assume no adjustment costs in the operation of production technology, as well as no costs of switching between occupations. In terms of entry (exit) into (from) entrepreneurship, the only cost is foregone wage income (profits). These assumptions render much of the problem static and make the solution considerably simpler, since each agent's occupational choice is a control variable in every period but not a state variable in the next one.

After occupational choices have been made the flow budget constraint becomes,

$$\dot{a} = \mathcal{Y}(h, a, z) + ra - c \tag{13}$$

$$\mathcal{Y}(h, a, z; w, r) = \max\left\{\pi(h, a, z; w, r), wh\right\} \tag{14}$$

$$a \geq 0 \tag{15}$$

IV.4. Recursive Formulation and Stationary Competitive Equilibrium

I restrict my attention to stationary competitive equilibria, i.e., $\partial V/\partial t = 0$. In recursive form, the set of optimization problems can be summarized by pairs of Hamilton-Jacobi-Bellman equations and a set of boundary conditions over the state space Ω,

$$\begin{aligned}
\rho V(h, a, z) = \max_{c\in\mathcal{C}}\ u(c) &+ V_a(h, a, z)\left(\mathcal{Y}(a, z, h; w, r) + ra - c\right)\\
&+ V_z(h, a, z)\widetilde{\mu}_z(z) + \frac{1}{2}V_{zz}\,\widetilde{\sigma}^2_z(z)\\
&+ V_h(h, a, z)\widetilde{\mu}_h(h) + \frac{1}{2}V_{hh}\,\widetilde{\sigma}^2_h(h)\\
= \max_{c\in\mathcal{C}}\ & u(c) + \mathcal{A}V(h, a, z)
\end{aligned} \tag{16}$$

$$V_a(h, \underline{a}, z) \geq u'\left(\mathcal{Y}(h, \underline{a}, z)\right) \quad \forall z, h \qquad (\text{SCBC for } \underline{a} = 0) \tag{17}$$

$$V_z(h, a, \underline{z}) = V_z(h, a, \bar{z}) = 0 \quad \forall a, h \qquad (\text{NBC for } \underline{z} \text{ and } \bar{z}) \tag{18}$$

$$V_h(\underline{h}, a, z) = V_h(\bar{h}, a, z) = 0 \quad \forall a, z \qquad (\text{NBC for } \underline{h} \text{ and } \bar{h}) \tag{19}$$

$$\lim_{t\to\infty} e^{-\rho t} a_t\, V_a(h, a, z) = 0 \qquad (\text{TVC}) \tag{20}$$

In a stationary equilibrium the state of the economy is given by the non-singular joint distribution $G(h, a, z)$, with corresponding density $g(h, a, z)$[13], which satisfies the

[13] As a side note, most papers either assume or impose the existence and uniqueness of the corresponding density, but it is in fact straightforward to prove it in a fairly general setting. Assuming the standard Lipschitz and Hólder conditions that ensure the existence and uniqueness of a SDE solution for each univariate Itô process $X_i, i = 1, \ldots, n$, it is known that the induced measure is absolutely continuous wrt the Lebesgue measure λ. By the absolute continuity of every induced measure μ_i, and since these measures are σ-finite as each state space is generated by sets of finite measure, it follows that the product measure $\mu^n := \mu_1 \times \cdots \times \mu_n$ is absolutely continuous wrt to the n-dimensional Lebesgue mea-

Kolmogorov Forward (or Fokker-Planck) equation,

$$\begin{aligned}
0 = &- \frac{\partial}{\partial a}\left[\widetilde{s}(h,a,z)\, g(h,a,z)\right] \\
&- \frac{\partial}{\partial z}\left[\widetilde{\mu}_z(z)\, g(h,a,z)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\widetilde{\sigma}_z^2(z)\, g(h,a,z)\right] \\
&- \frac{\partial}{\partial h}\left[\widetilde{\mu}_h(h)\, g(h,a,z)\right] + \frac{1}{2}\frac{\partial^2}{\partial h^2}\left[\widetilde{\sigma}_h^2(h)\, g(h,a,z)\right] \\
&= \mathcal{B}\, g(h,a,z)
\end{aligned} \tag{21}$$

with the necessary boundary condition,

$$0 = -\Big[v\,(\widetilde{s}\, g)\Big]_{\mathcal{A}} + \left[v\left(-\widetilde{\mu}_z\, g + \frac{1}{2}\frac{\partial}{\partial z}(\widetilde{\sigma}_z^2\, g)\right)\right]_{\mathcal{Z}} + \left[v\left(-\widetilde{\mu}_h\, g + \frac{1}{2}\frac{\partial}{\partial h}(\widetilde{\sigma}_h^2\, g)\right)\right]_{\mathcal{H}} \tag{22}$$

The term $\widetilde{s}(h,a,z) := (\mathcal{Y}(a,z,h;w,r) + ra - \widetilde{c})$ denotes optimal savings, i.e., the optimal drift that solves the corresponding HJB equation. Note that the "HJB operator", $\mathcal{A}$, is the infinitesimal generator of the joint stochastic process (h,a,z). Such a differential operator can be thought of as the infinite-dimensional analogue of a continuous time transition matrix. Accordingly, the "KFE operator" $\mathcal{B}$ has a very special connection with the HJB equation.

LEMMA 3: *The Kolmogorov Forward differential operator $\mathcal{B}$ in (21), with the necessary boundary condition (22), is the (Hermitian) adjoint of the Hamilton-Jacobi-Bellman infinitesimal generator $\mathcal{A}$ in (16) with boundary conditions (17) through (19). That is, $\mathcal{B} = \mathcal{A}^*$.*

A stationary competitive equilibrium consists of a set of $\mathcal{F}_t$-adapted allocation processes $\{\widetilde{c}_t, \widetilde{s}_t, \widehat{\ell}_t, \widehat{k}_t, \mathcal{Y}_t\}_{t=0}^{\infty}$, price processes $\{w_t, r_t\}_{t=0}^{\infty}$, and sequences of joint distributions $\{G(h,a,z)\}_{t=1}^{\infty}$, such that, for all $t \geq 0$:

1) Given price processes, the allocations $\widetilde{c}_t(h,a,z), \widetilde{s}_t(h,a,z), \widehat{\ell}_t(h,a,z), \widehat{k}_t(h,a,z), \mathcal{Y}_t(h,a,z)\}$ solve the stochastic control problems (16)–(19)

2) The physical capital/asset market clears:

$$S(r,w) := \iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} a\; dG(h,a,z) = \iiint\limits_{\mathcal{H}_{\mathcal{E}}\times\mathcal{A}_{\mathcal{E}}\times\mathcal{Z}_{\mathcal{E}}} \widehat{k}(h,a,z)\; dG(h,a,z) =: K(r,w)$$

sure λ^n. As $\mu^n << \lambda^n$, the Radon-Nikodym theorem ensures the existence, uniqueness (up to λ^n-null sets), and non-negativity λ^n-almost everywhere of a joint density $g(x_1,\ldots,x_n) \geq 0$, which is simply the Radon-Nikodym derivative of the induced product measure wrt λ^n.

3) The human capital/labor market clears:

$$H(r,w) := \iiint\limits_{\mathcal{H}_\mathcal{W}\times\mathcal{A}_\mathcal{W}\times\mathcal{Z}_\mathcal{W}} h \; dG(h,a,z) = \iiint\limits_{\mathcal{H}_\mathcal{E}\times\mathcal{A}_\mathcal{E}\times\mathcal{Z}_\mathcal{E}} \widehat{\ell}(h,a,z)\; dG(h,a,z) =: L(r,w)$$

4) The goods market clears (by Walras's law):

$$Y(r,w) := \iiint\limits_{\mathcal{H}_\mathcal{E}\times\mathcal{A}_\mathcal{E}\times\mathcal{E}} \widehat{y}(h,a,z)\; dG(h,a,z) = \iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} \widetilde{c}(h,a,z)\; dG(h,a,z) + \delta K(r,w)$$

5) The joint density of human capital, wealth, and entrepreneurial ability satisfies the Kolmogorov-Forward equation (21) with boundary condition (22).

Obtaining an equilibrium amounts to solving pairs of coupled, nonlinear, second-order partial differential equations (HJB and KFE). Classical PDE solutions are not guaranteed to hold in this setting, as the non-convexities induced by occupational and technology choices give rise to convex kinks. The appropriate solution method requires the powerful theory of constrained *viscosity solutions*; see Crandall and Lions (1983).

The assumption of a continuum of non-atomic agents that face idiosyncratic – but not aggregate shocks – and interact strategically only through the incentives actuated by a common set of prices (r, w), implies the following result. Given solutions to HJB equations and the market clearing conditions, despite the stochastic evolution of individual state variables, the dynamics of the aggregate state distribution $G(h, a, z)$ is deterministic due to a law of large numbers with its density governed by the Kolmogorov Forward equation. This is a multivariate case of what Lasry and Lions (2007) call a *Mean Field Game* without common noise. As a result, joint/conditional/marginal frequency and probability distributions coincide on all subsets of the unit interval, hence we can readily obtain aggregate quantities through multiple integration over Ω.[14]

Computationally, the problem is solved using an implicit upwind finite difference scheme that is stable, consistent, and monotone in the sense of Barles and Souganidis (1991), and thus yields unique viscosity solutions to the HJB and KF equations. For a detailed exposition see the work Achdou et al. (2021), the approach of whom I follow closely. All computational details, including the construction of grids for (h, a, z) and the solution algorithm, are relegated to Appendix B.

[14] To avoid any potential measurability problems, one can appeal to the theorems of Uhlig (1996). Moreover, in this setting we can always switch from multiple to repeated integration by the Fubini-Tonelli theorem: given the σ-finiteness of measure spaces and the measurability of functions, the iterated integrals are well-defined with respect to the unique product measure and equal to the multiple integrals.

As in many studies in the literature, an object of interest is the productivity cutoff above which individuals decide to become entrepreneurs. In contrast to most studies, this threshold is neither unique for all agents, nor does it hinge on wealth alone, but also depends on the level of human capital in a non-linear way. Each cutoff is jointly determined by a three-dimensional fictitious marginal agent that is indifferent between becoming an entrepreneur or a wage worker. A first finding is summarized below.

LEMMA 2 (NON-MONOTONE PRODUCTIVITY CUTOFFS): *(a) The entrepreneurial productivity cutoff for unconstrained agents ($\widehat{r} = r$) is independent of assets and given by:*

$$\underline{z}(h_e) = \begin{cases} h_e^{\nu} \left(\frac{w}{\nu}\right)^{\nu} \left[(1-\nu)\left(\frac{\alpha}{r+\delta}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\nu-1} & \text{for } h_e \leq \hbar \\ h_e^{\nu-\omega}\, \hbar^{\omega} \left(\frac{w}{\nu}\right)^{\nu} \left[(1-\nu)\left(\frac{\alpha}{r+\delta}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\nu-1} & \text{for } h_e > \hbar \end{cases} \tag{23}$$

Therefore, $\underline{z}(h_e)$ is strictly increasing in human capital for all $h_e \leq \hbar$, and under the condition $\omega > \nu$, it is strictly decreasing for all $h_e > \hbar$.

(b) The entrepreneurial productivity cutoff for each constrained agent i depends on her assets, $a_i < (1-\vartheta)k(a_i, z, h_e)$, and given by:

$$\underline{z}^c(h_e, a_i) = \begin{cases} h_e^{1-\upsilon} \left(\frac{w}{1-\upsilon}\right)^{\frac{1}{1-\upsilon}} \left[(\lambda a_i)^{\alpha(1-\nu)}\left(\frac{\upsilon}{w}\right)^{\upsilon} - (\lambda a_i)(r+\delta)\right]^{\upsilon-1} & \text{for } h_e < \hbar \\ h_e^{1-\upsilon-\omega}\, \hbar^{\omega} \left(\frac{w}{1-\upsilon}\right)^{\frac{1}{1-\upsilon}} \left[(\lambda a_i)^{\alpha(1-\nu)}\left(\frac{\upsilon}{w}\right)^{\upsilon} - (\lambda a_i)(r+\delta)\right]^{\upsilon-1} & \text{for } h_e \geq \hbar \end{cases} \tag{24}$$

where $\upsilon := (1-\alpha)(1-\nu)$. The cutoff $\underline{z}^c(h_e, a)$ is strictly increasing in human capital for all $h_e \leq \hbar$, and under the condition $\omega > \alpha + \nu - \alpha\nu$, it is strictly decreasing for all $h_e > \hbar$.

Lemma 2 demonstrates that occupational choices depend on human capital non-monotonically, under a necessary and sufficient condition. As we consider individuals with human capital less than $\hbar$, breaking even requires draws from a progressively higher part of the entrepreneurial talent distribution, which naturally occurs with lower probability. On the contrary, higher human capital when $h > \hbar$ makes the entry into entrepreneurship progressively easier (in terms of z-cutoffs) since profits grow faster than labor income in that direction of h, as long as the complementarity parameter ω is not too low. These results also show that the tightness of the collateral constraint will impact the occupational choice of individuals who would be constrained should they decided to become entrepreneurs. For a given wealth level, tighter borrowing conditions make it more difficult for profits to match foregone labor income, inducing a higher cutoff for z.

PROPOSITION 1: For any sequence of equilibrium prices $\{w_t, r_t\}_{t\in\mathbb{R}_+}$, joint distributions $\{G_t(h, a, z)\}_{t\in\mathbb{R}_+}$, technology adoption cost $\kappa \in [0, 1]$, and collateral constraint $\vartheta \in [0, 1]$, there exists a non-increasing measurable function $\chi(\vartheta) : [0, 1] \mapsto [\nu, \alpha + \nu - \alpha\nu]$, such that the entrepreneurship rate by human capital is:

- strictly decreasing $\forall h \in \mathcal{H}$ *almost surely*, if and only $\omega < \chi(\vartheta)$;
- strictly decreasing $\forall h < \hbar(\kappa, \omega)$ and strictly increasing $\forall h \geq \hbar(\kappa, \omega)$ *almost surely*, if and only if $\omega > \chi(\vartheta)$;
- strictly decreasing $\forall h < \hbar(\kappa, \omega)$ *almost surely*, and anything goes $\forall h \geq \hbar(\kappa, \omega)$, if and only if $\nu < \omega \leq \chi(\vartheta)$.

The conclusions of Proposition 1 are stark, flexible, and in line with the empirical findings under minimal restrictive assumptions. The pivotal condition bears on the complementarity parameter in tandem with financial frictions; it requires ω to be larger than an endogenous cutoff value that is non-decreasing in ϑ, meaning that tighter financial markets dictate a larger ω to achieve entrepreneurship rates that are increasing somewhere in the support of human capital. These results hold under *any* joint distribution for (z, h) that is non-flat almost everywhere in its support. The steepness and curvature of this relationship will obviously depend on the magnitude of parameters, but the main consequences are nonetheless clear-cut. For example, the probability of being an active entrepreneur in this economy will arise as asymmetric U-shaped in the second case the proposition, with the lowest probability occurring right on the threshold $\hbar(\kappa, \omega)$.

This global minimum property offers not only a sharp theoretical prediction, but also a decisive empirical advantage. Given a mapping between years of schooling and human capital, $\hbar$ is an identifiable parameter determined by observables – either directly from descriptive data or through a reduced-form model – that can be calibrated or estimated using standard structural methods. This makes the model particularly attractive to use with micro-level datasets, as well as for cross-country analysis.

PROPOSITION 2: *For any sequence of equilibrium prices $\{w_t, r_t\}_{t\in\mathbb{R}_+}$, joint distributions $\{G_t(h, a, z)\}_{t\in\mathbb{R}_+}$, technology adoption cost $\kappa \in [0, 1]$, and collateral constraint $\vartheta \in [0, 1]$, at each time t, the expected value of any entrepreneurial outcome X (production plans, profits, capital and labor demands) is strictly increasing in human capital with probability 1. That is, $\frac{\partial \mathbb{E}[X|\mathcal{E}, h_e]}{\partial h_e} > 0$ a.s. $\forall h \in \mathcal{H}$.*

Proposition 2 establishes another key result that is prevalent in the empirics of Section III: the central tendency of business outcomes is monotonically increasing in the educational attainment of the firm owner/entrepreneur. The result is guaranteed to hold irregardless of the value of ω, and the level of the cost κ is extraneous to the core prediction. This occurs for two reasons. For $h < \hbar$, there is a self-selection effect: choosing entrepreneurship instead of working for a wage implies higher draws from the z-distribution as h increases in order to . For $h \geq \hbar$, technological choices lead to a rightward shift of the effective productivity distribution; i.e., for the same level of z, the "more educated" distribution dominates the "less educated" one in the first-order stochastic dominance sense.

PROPOSITION 3 (AGGREGATION): *In a competitive equilibrium, the production side of the economy aggregates and net total output per person can be expressed as:*

$$Y = \frac{\left(\iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} \left(z\widetilde{h}\right)^{\frac{1}{\nu}} \left(\widehat{r}(h,a,z)+\delta\right)^{-\frac{\widehat{\alpha}}{\nu}} dG(h,a,z,\mathcal{E}) \right)^{\widehat{\alpha}+\nu}}{\left(\iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} \left(z\widetilde{h}\right)^{\frac{1}{\nu}} \left(\widehat{r}(h,a,z)+\delta\right)^{-\frac{(\widehat{\alpha}+\nu)}{\nu}} dG(h,a,z,\mathcal{E}) \right)^{\widehat{\alpha}}} \left(K^{\alpha}H^{1-\alpha}\right)^{1-\nu}$$

$$= \underbrace{\mu(\mathcal{E})^{\nu}}_{\text{TFP extensive margin}} \underbrace{\frac{\left(\mathbb{E}\left[\left(z\widetilde{h}\right)^{\frac{1}{\nu}} \left(\widehat{r}(h,a,z)+\delta\right)^{-\frac{\widehat{\alpha}}{\nu}} \,\middle|\, \mathcal{E} \right] \right)^{\widehat{\alpha}+\nu}}{\left(\mathbb{E}\left[\left(z\widetilde{h}\right)^{\frac{1}{\nu}} \left(\widehat{r}(h,a,z)+\delta\right)^{-\frac{(\widehat{\alpha}+\nu)}{\nu}} \,\middle|\, \mathcal{E} \right] \right)^{\widehat{\alpha}}}}_{\text{TFP intensive margin}} \left(K^{\alpha}H^{1-\alpha}\right)^{1-\nu} \tag{25}$$

where $\widehat{\alpha} := \alpha(1-\nu)$ and $\mu(\mathcal{E})$ is the measure of entrepreneurs. In the absence of financial frictions $(\vartheta \to 1)$, total output per worker becomes:

$$Y_* = \mu(\mathcal{E}_*)^{\nu} \left(\mathbb{E}\left[\left(z\widetilde{h}\right)^{\frac{1}{\nu}} \,\middle|\, \mathcal{E}_* \right] \right)^{\nu} \left(K_*^{\alpha}H_*^{1-\alpha}\right)^{1-\nu} \tag{26}$$

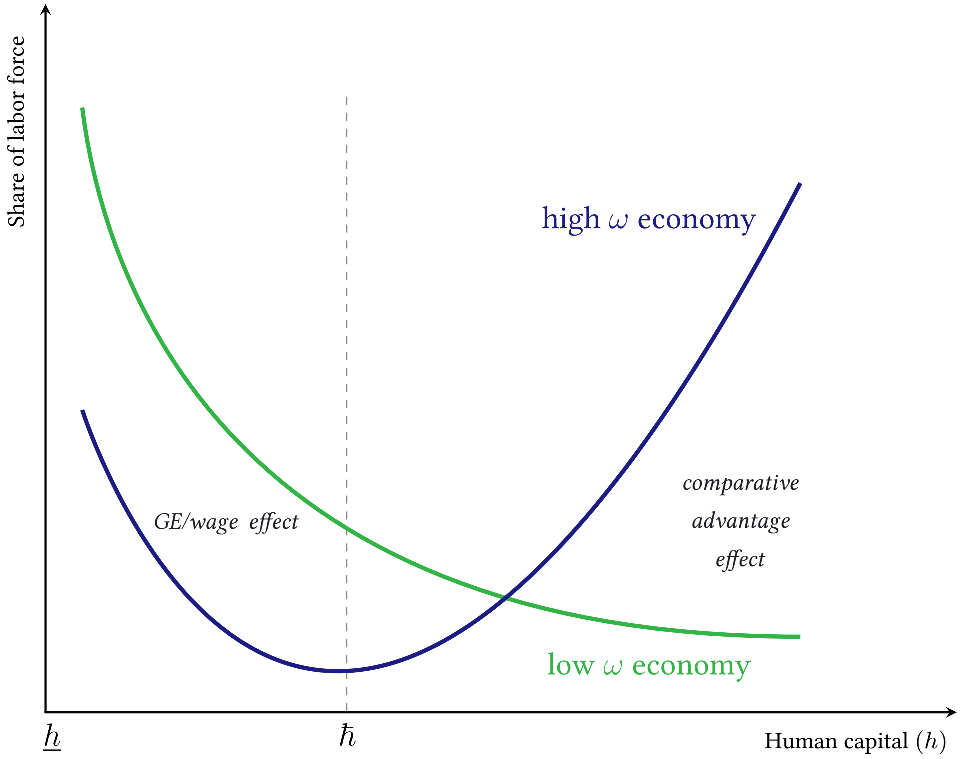
To sun up, it would be instructive to use a simple illustrative diagram to contrast the implications of the model economy .

V. Quantitative Exploration

V.1. Parametrization and Calibration

The length of a period is taken to be one year to allow for internal consistency with aggregate, survey, and firm-level data. The calibration strategy is on the side of par-

Figure 6: Entrepreneurship Rate along h: High vs Low Complementarity



Notes: This illustrative diagram depicts the fraction of the labor force opting into entrepreneurship for each level of human capital h. The curve for "this economy" is drawn for the empirically relevant case of $\omega > \omega(\vartheta) \in [\alpha+\nu-\alpha\nu]$. The term "standard economies" refers to macro models that do not allow for human capital/labor ability to directly affect entrepreneurial outcomes, or do not impose any functional relationship between the distributions of z and h.

simony and contrary to the proliferation of free parameters. I reduce the number of degrees of freedom first by preassigning values to conventional and well-accepted parameters in the literature, next by estimating parameters that can be easily inferred from the data used in the empirical section of the paper. The model is then disciplined via the joint calibration of only six parameters that are relatively nonstandard, and thus of principal interest. Table 3 summarizes the results discussed below.

To begin with, we need to specify the exact type of Itô processes for z_t and h_t. One straightforward and reasonable option is to model their evolution by,

$$d \log z_t = \varphi_z(\mu_z - \log z_t)\, dt + \sigma_z\, dW_t^z \tag{27}$$

$$d \log h_t = \varphi_h(\mu_h - \log h_t)\, dt + \sigma_h\, dW_t^h \tag{28}$$

Each equation describes a Ornstein-Uhlenbeck (OU) process in natural logs – the continuous-time analogue of a log AR(1) process, with drift μ, speed of reversion (persistence) governed by φ, and innovation dispersion σ. An application of Itô's lemma yields the representation of each process in levels,

$$dz_t = \left[\varphi_z(\mu_z - \log z_t) + \frac{1}{2}\sigma_z^2\right] z_t\, dt + \sigma_z z_t\, dW_t^z \tag{29}$$

$$dh_t = \left[\varphi_h(\mu_h - \log h_t) + \frac{1}{2}\sigma_h^2\right] h_t\, dt + \sigma_h h_t\, dW_t^h \tag{30}$$

By virtue of mean-reversion and the properties of W_t, both (marginal) stationary and ergodic distributions are log-normal: $\log z_t \sim \mathcal{N}(\mu_z, \frac{\sigma_z^2}{2\varphi_z})$ and $\log h_t \sim \mathcal{N}(\mu_h, \frac{\sigma_h^2}{2\varphi_h})$. By the assumption of independent Wiener processes, (z_t, h_t) are also jointly log-normal at the population level.[15] Since the distribution of shocks to entrepreneurial productivity predominantly shapes the firm size distribution in the model, the above result is rather empirically relevant; see the work of Kondo, Lewis and Stella (2021).[16]

The model requires 14 parameter values to be specified. On the household side, the subjective discount ρ and coefficient of relative risk aversion γ; two technological parameters, α and ν; the capital depreciation rate δ; six parameters describing the OU processes for entrepreneurial productivity, $\mu_z, \sigma_z, \varphi_z$, and human capital, $\mu_h, \sigma_h, \varphi_h$; the extent of financial frictions captured by ϑ; and the level of the local threshold externality $\hbar$ together with the complementarity parameter ω.

Preassigned parameters: The set of preassigned parameters is $\{\alpha, \gamma, \delta, \mu_z, \varphi_h\}$. Values for the first three are fairly canonical in the macro-development and growth literature. I fix the coefficient of relative risk aversion to $\gamma = 1.5$; the annual capital depreciation rate is set to $\delta = 0.06$; the elasticity of output with respect to capital is chosen to be $\alpha = 0.36$, which, in the presence of moderate financial frictions, results in an aggregate capital share of income close to 0.34 for the U.S.. As normalization, the drift of entrepreneurial productivity is $\mu_z = 0$, which simply implies a zero-mean process in logs. Lastly I set

[15] Of course this does not imply uncorrelatedness or joint normality *within* each occupational group. Even with frictionless capital markets, productivity cutoffs still depend on human capital, thus the joint distribution of (z, h) for entrepreneurs (workers) will turn out to have some dependency structure.

[16] Using an extensive confidential Census Bureau panel dataset, one of the main stylized facts the authors establish is that "a lognormal fits both firm and establishment size employment distributions better than the commonly used Pareto, even far in the truncated upper tail." This is an important empirical finding that casts doubt on the conventional wisdom that the U.S. firm size distributions are adequately approximated by a Pareto distribution with a shape parameter close to one (Zipf's law).

the persistence parameter for h_t to the highest value possible without causing numerical problems, $e^{-\varphi_h} = 0.999$, for two reasons. First, since I am mapping the model notion of human capital to educational attainment in the data, I am identifying a characteristic that is practically time-invariant. Second, I want to abstain from sizeable shocks to human capital so that the main source of firm dynamics is entrepreneurial productivity – and secondarily consumption-savings dynamics – thus making my results more easily comparable to the majority of the literature.

Estimated parameters: There are three parameters that can be easily estimated using the data and methods of Section III: $\{\mu_h, \sigma_h, \hbar\}$. These are all related to human capital, so as long as we specify a bijective map between units of human capital and measured years of formal education, statistical inference is straightforward. Regarding the former, I draw upon a standard practice in the macro-development literature, e.g., Hall and Jones (1999), Caselli (2005), and assume the simple mapping,

$$h_i = e^{\theta(S_i)}, \quad \theta(S) = \begin{cases} 0.134 \cdot S & \text{if } S \leq 4 \\ 0.134 \cdot 4 + 0.101 \cdot (S-4) & \text{if } 4 < S \leq 8 \\ 0.134 \cdot 4 + 0.101 \cdot 4 + 0.068 \cdot (S-8) & \text{if } 8 < S \end{cases} \tag{31}$$

where $S = 0, 1, 2, \ldots,$ denotes years of formal education and $\theta(S)$ is a piecewise-linear function that reconciles the Mincerian approach (log-linearity of wages and schooling) at the country level with the concavity of this relationship across countries, e.g., Psacharopoulos (1994).

Using this transformation I estimate the parameters of a log-normal distribution with location μ_h and scale $s_h^2 = \sigma_h^2/(2\varphi_h)$ using MLE in the pooled SCF data for $1989 - 2019$, resulting in $\log h \sim \mathcal{N}(1.3421, 0.0502)$; equivalently, $\sigma_h = 0.010$. Finally, I employ the reduced-form results of Section III to identify the level of the threshold externality. Depending on how many controls are considered, the estimated relationship ranges from the 8th year to the 10th year of schooling. An evenhanded option is to choose the 9th year of schooling.

Calibrated parameters: There remain six parameters to be jointly calibrated for the model to best fit six relevant moments in the data. The vector under consideration is $\{\nu, \omega, \sigma_z, \varphi_z, \vartheta, \rho\}$. Although exact identification is typically impossible in this class of non-linear general equilibrium models, the targeted moments are selected to be sufficiently informative about the calibrated parameters so that the objective function is not flat along any direction. The calibration sensitivity matrix confirms and validates

Table 3 reports the output of the calibration exercise and summarizes the rest of the parametrization. Many of the calibrated parameter values are easily comparable and are close to alternative estimates of these values in the relevant literature.

$\{\nu \models$ *fraction of entrepreneurs in the labor force*$\}$. The span-of-control parameter governs the shape and scale of production possibilities and profit functions, hence is of paramount significance when in comes to occupational decisions. The same conclusion can be reached by observing the dominant role of ν in the z-cutoff expressions of Lemma 2. The calibrated value is $\nu = 0.232$. At first glance this may seem higher compared to economies without financial frictions and/or with a corporate sector, typically in the range of $0.15 - 0.20$, but it is not so in comparison to the germane class of models. For instance, the calibration of Buera, Kaboski and Shin (2011) to the U.S. economy results in $\nu = 0.21$ ($\alpha + \theta = 0.79$ in their model), which in turn generates an entrepreneurship rate of only 5 percent.

$\{\sigma_z \models$ *dispersion of (log) TFP*$\}$ The standard deviation of entrepreneurial productivity shocks is clearly the primary source of variation among production plans, it is thus palpable to match the cross-sectional dispersion of firms' log TFP in the data. A good benchmark choice is given by Haltiwanger (2011), who states that the bulk of empirical evidence suggests "…estimates of the standard deviation of innovations to productivity shocks of about 0.20 (in terms of log total factor productivity)." This observation is also in line with the summary statistics on firms' productivity measures in Foster, Haltiwanger and Syverson (2008).

$\{\varphi_z \models$ *firm entry rate*$\}$ The persistence of entrepreneurial productivity is by far the main driver of firm dynamics in this setting, and so it is chosen to match the average U.S. firm entry rate for $1989-2019$. As a reminder, in the stationary equilibrium of the model economy the firm entry and exit rate are equal. The calibrated value of $e^{-\varphi_z} = 0.908$ is close to the estimates of Asker, Collard-Wexler and De Loecker (2014) using an AR(1) specification for (the log of) TFPR. In fact, it falls between their own estimates and the reproduction of the Hsieh and Klenow (2009) analysis using their dataset. [17]

$\{\vartheta \models$ *External finance-to-output ratio*$\}$ the most relevant data counterpart is at least since the work of Rajan and Zingales (1998) [18]

$\{\rho \models$ *real rate of return on capital*$\}$ The household discount rate, ρ, is set to match an average real interest rate of 4 percent per annum, a value commonly used in the literature. In view of r as the average real rate of return on capital, it also conforms with findings by Gomme, Ravikumar and Rupert (2011). The calibrated value of $\rho = 0.089$ implies an economy with fairly impatient agents, a result that is fairly standard in the sphere of incomplete-market models with entrepreneurship. For instance, the baseline calibration of Cagetti and De Nardi (2006) requires a (discrete time) discount factor of $\beta = 0.865$ to match a 6.5 percent equilibrium interest rate; similarly the model of Buera and Shin (2013) targets a 4.5 percent interest rate and requires a discount factor of $\beta = 0.904$.

[17] See Table 1 at http://www.johnasker.com/ACWDLcomment.pdf.

[18] See for example Buera, Kaboski and Shin (2011), Buera and Shin (2013), Midrigan and Xu (2014), Moll (2014) and the references therein.

Table 3: Baseline Model Calibration; Moments and Parameters

Targeted Moments	Data	Model	Source
Entrepreneurship rate in the labor force	0.105	0.105	SCF and CPS
Share of total income received by top 20%	0.499	0.492	CPS ASEC
Firm entry/exit rate	0.110	0.110	BDS
Dispersion of (log) TFP	0.200	0.200	Haltiwanger (2011)
External finance-to-output ratio	1.640	1.640	GFDD/World Bank
Real rate of return on capital	0.040	0.040	See text
Calibrated Parameters	**Value**		**Comment**
Span-of-control parameter (ν)	0.232		xxx
Complementarity between z and h_e (ω)	0.511		xxx
Dispersion of productivity shock (σ_z)	0.210		xxx
Autocorrelation of productivity z ($e^{-\varphi_z}$)	0.908		xxx
Collateral constraint (ϑ)	0.866		xxx
Subjective discount rate (ρ)	0.089		xxx
Preassigned Parameters	**Value**		**Comment**
Elasticity of output w.r.t. capital (α)	0.360		Standard
Capital depreciation rate (δ)	0.060		Standard
Coefficient of relative risk aversion (γ)	1.500		Standard
Mean (log) entrepreneurial productivity (μ_z)	0.000		Normalization
Autocorrelation of human capital ($e^{-\varphi_h}$)	0.999		xxx
Estimated Parameters	**Value**		**Comment**
Mean (log) human capital (μ_h)	1.342		MLE based on the grid for h
Dispersion of human capital shock (σ_h)	0.010		MLE based on the grid for h
Threshold externality ($\hbar$)	9 years		Probit regressions in Section III

Table 4: Baseline Model Output; Untargeted Moments

Untargeted Moments	Data	Model	Source
Fraction of wealth owned by entrepreneurs	0.437	0.441	SCF
Ratio of mean assets; entrepreneur to worker	6.876	6.753	SCF
Share of total income received by the top quintile	0.499	0.491	CPS ASEC
Share of total income received by the middle quintile	0.148	0.134	CPS ASEC
Job destruction rate: deaths	0.047	0.041	BDS
Fraction of entrepreneurs with $16+$ years of education	0.438	0.438	SCF
Fraction of entrepreneurs with $12-$ years education	0.329	0.329	SCF
Firm size distribution:			
Employment size: $1-9$	0.769	0.775	BDS
Employment size: $10-19$	0.115	0.129	BDS
Employment size: $20-99$	0.096	0.093	BDS
Employment size: $100+$	0.020	0.003	BDS
Entrepreneurship rate by educational attainment:			
Years of education: $[0-8]$	0.080	0.080	SCF
Years of education: $(8-12)$	0.072	0.072	SCF
Years of education: 12	0.091	0.091	SCF
Years of education: $(12-16)$	0.094	0.099	SCF
Years of education: 16	0.121	0.121	SCF
Years of education: $(16, 21]$	0.156	0.156	SCF

REFERENCES

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2021. "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach." *The Review of Economic Studies*, 89(1): 45–86.

Aghion, Philippe, and Patrick Bolton. 1997. "A Theory of Trickle-Down Growth and Development." *Review of Economic Studies*, 64(2): 151–172.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy*, 122(5): 1013–1063.

Banerjee, Abhijit, and Benjamin Moll. 2010. "Why Does Misallocation Persist?" *American Economic Journal: Macroeconomics*, 2(1): 189–206.

Banerjee, Abhijit V., and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy*, 101(2): 274–298.

Barles, Guy, and Panagiotis E. Souganidis. 1991. "Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations." *Asymptotic Analysis*, 4(3): 271–283.

Barnard, J, and DB Rubin. 1999. "Small-sample degrees of freedom with multiple imputation." *Biometrika*, 86(4): 948–955.

Blanchflower, David. 2000. "Self-employment in OECD countries." *Labour Economics*, 7(5): 471–505.

Blanchflower, David G. 2004. "Self-Employment: More may not be better." National Bureau of Economic Research Working Paper 10286.

Buera, Francisco, and Yongseok Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy*, 121(2): 221 – 272.

Buera, Francisco, Joseph Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review*, 101(5): 1964–2002.

Buera, Francisco, Joseph Kaboski, and Yongseok Shin. 2015. "Entrepreneurship and Financial Frictions: A Macrodevelopment Perspective." *Annual Review of Economics*, 7(1): 409–436.

Cagetti, Marco, and Mariacristina De Nardi. 2006. "Entrepreneurship, Frictions, and Wealth." *Journal of Political Economy*, 114(5): 835–870.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth*. Vol. 1 of *Handbook of Economic Growth*, , ed. Philippe Aghion and Steven Durlauf, Chapter 9, 679–741. Elsevier.

Crandall, Michael G., and Pierre-Louis Lions. 1983. "Viscosity Solutions of Hamilton-Jacobi Equations." *Transactions of the American Mathematical Society*, 277(1): 1–42.

De Nardi, Mariacristina, Phil Doctor, and Spencer Krane. 2007. "Evidence on entrepreneurs in the United States: data from the 1989-2004 survey of consumer finances." *Economic Perspectives*, 31(Q IV): 18–36.

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

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *The American Economic Review*, 98(1): 394–425.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The Dynamics of Inequality." *Econometrica*, 84(6): 2071–2111.

Gomme, Paul, B Ravikumar, and Peter Rupert. 2011. "The Return to Capital and the Business Cycle." *Review of Economic Dynamics*, 14(2): 262–278.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *The Quarterly Journal of Economics*, 114(1): 83–116.

Haltiwanger, John. 2011. " Firm Dynamics and Productivity Growth." *European Investment Bank Papers*, 16(1): 116–136.

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

Hipple, Steven F. 2010. "Self-Employment in the United States." *Monthly Labor Review*, 133(9): 17–32.

Hsieh, Chang-Tai, and Pete Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Kennickell, Arthur B. 2017. "Multiple Imputation in the Survey of Consumer Finances." *Statistical Journal of the IAOS*, 33(1): 143–151.

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

Kitao, Sagiri. 2008. "Entrepreneurship, Taxation and Capital Investment." *Review of Economic Dynamics*, 11(1): 44–69.

Knight, Frank H. 1921. *Risk, Uncertainty and Profit.* Hart, Schaffner, and Marx; Houghton Mifflin.

Kondo, Illenin O., Logan T. Lewis, and Andrea Stella. 2021. "Heavy Tailed, but not Zipf: Firm and Establishment Size in the U.S." Center for Economic Studies, U.S. Census Bureau Working Papers 21-15.

Lasry, Jean-Michel, and Pierre-Louis Lions. 2007. "Mean Field Games." *Japanese Journal of Mathematics*, 2: 229–260.

Little, Roderick J. A. 1992. "Regression with Missing X's: A Review." *Journal of the American Statistical Association*, 87(420): 1227–1237.

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

Midrigan, Virgiliu, and Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review*, 104(2): 422–58.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *American Economic Review*, 104(10): 3186–3221.

Piketty, Thomas. 1997. "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing." *The Review of Economic Studies*, 64(2): 173–189.

Poschke, Markus. 2013. "Who becomes an entrepreneur? Labor market prospects and occupational choice." *Journal of Economic Dynamics and Control*, 37(3): 693 – 710.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." *World Development*, 22(9): 1325–1343.

Quadrini, Vincenzo. 2009. "Entrepreneurship in Macroeconomics." *Annals of Finance*, 5(3): 295–311.

Rajan, Raghuram G, and Luigi Zingales. 1998. "Financial Dependence and Growth." *American Economic Review*, 88(3): 559–586.

Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics*, 311–323.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys.* John Wiley and Sons.

Schumpeter, J.A. 1911. *Theorie der Wirtschaftlichen Entwicklung (The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle).* Leipzig, Germany: Duncker und Humblot.

Tavella, Domingo, and Curt Randall. 2000. *Pricing Financial Instruments: The Finite Difference Method.* John Wiley and Sons, New York.

Uhlig, Harald. 1996. "A Law of Large Numbers for Large Economies." *Economic Theory*, 8(1): 41–50.

Van Buuren, Stef. 2018. *Flexible Imputation of Missing Data, Second Edition.* New York: Chapman and Hall/CRC.

A1. Proofs and Derivations

Proof of Lemma 2**:** We start by deriving entrepreneurial factor demands. In the unconstrained case when the individual's wealth is $a > \widehat{k}(a, z, h_e|r, w)/\lambda$, the demand for capital (k_*) and labor (l_*) comes from a standard profit maximization exercise.

$$\{k_*,\, l_*\} = \arg\max_{\substack{k\geq 0\\ l\geq 0}} z\tilde{h}\left(k^{\alpha}l^{1-\alpha}\right)^{1-\nu} - wl - (r+\delta)k$$

The conditional demand for capital and labor come from rearranging the relevant FOCs,

$$k_* = \frac{\alpha\,(1-\nu)}{R}\,y_* \tag{A1}$$

$$l_* = \frac{(1-\alpha)(1-\nu)}{w}\,y_* \tag{A2}$$

where $R := r+\delta$. Substituting A1 and A2 into the production function and manipulating through yields the optimal level of production,

$$y_* = (z\tilde{h})^{\frac{1}{\nu}}\left[\left(\frac{\alpha(1-\nu)}{R}\right)^{\alpha}\left(\frac{(1-\alpha)(1-\nu)}{w}\right)^{1-\alpha}\right]^{\frac{1-\nu}{\nu}} \tag{A3}$$

We can easily recover unconstrained factor demands by plugging A3 into A1 and A2,

$$k_*(h_e, z|w, r) = (z\tilde{h})^{\frac{1}{\nu}}\left(\frac{\alpha(1-\nu)}{R}\right)^{\frac{1-(1-\alpha)(1-\nu)}{\nu}}\left(\frac{(1-\alpha)(1-\nu)}{w}\right)^{\frac{(1-\alpha)(1-\nu)}{\nu}} \tag{A4}$$

$$l_*(h_e, z|w, r) = (z\tilde{h})^{\frac{1}{\nu}}\left(\frac{\alpha(1-\nu)}{R}\right)^{\frac{\alpha(1-\nu)}{\nu}}\left(\frac{(1-\alpha)(1-\nu)}{w}\right)^{\frac{1-\alpha(1-\nu)}{\nu}} \tag{A5}$$

In general, under any degree of financial market imperfections, we can compactly express the demand for capital and labor as,

$$\widehat{k}(h_e, a, z|w, r) = \min\{\lambda a, k_*\} \tag{A6}$$

$$\widehat{l}(h_e, a, z|w, r) = \left((z\tilde{h})\,\frac{(1-\alpha)(1-\nu)}{w}\right)^{\frac{1}{1-(1-\alpha)(1-\nu)}}\widehat{k}(h_e, a, z|w, r)^{\frac{\alpha(1-\nu)}{1-(1-\alpha)(1-\nu)}} \tag{A7}$$

■

Proof of Lemma 2**:** Using the results of Lemma 1 it is straightforward to derive indirect profit functions. Maximized profits for unconstrained entrepreneurs are independent of their wealth and are given by,

$$\pi^{uc}(h_e, z \mid r, w) = (z\widetilde{h})^{\frac{1}{\nu}}\ \nu \left[(1-\nu)\left(\frac{\alpha}{r+\delta}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{1-\nu}{\nu}} \tag{A8}$$

On the other hand, maximized profits for constrained entrepreneurs depend on their wealth level and for $a < \frac{k(a,z,h_e|r,w)}{\lambda}$ are given by,

$$\pi^{c}(h_e, a, z \mid r, w) = (1-\upsilon)\left[(z\widetilde{h})\,(\lambda a)^{\alpha(1-\nu)}\left(\frac{\upsilon}{w}\right)^{\upsilon}\right]^{\frac{1}{1-\upsilon}} - (\lambda a)(r+\delta) \tag{A9}$$

where $\upsilon := (1-\alpha)(1-\nu)$. Setting up the equations $\pi^{uc}(h_e, z \mid r, w) = wh_e$ and $\pi^{c}(h_e, a, z \mid r, w) = wh_e$, using A8 and A9, and manipulating through to solve for z, yields the desired results. ■

To reduce burden on the reader I will use shorthand notation based on discretized probability distributions, with the understanding that the calculations involve proper integrals over specific domains, e.g., $\mathcal{P}[h = k]$ refers to $\int_{k_1}^{k_2} g(h)\,dh$. Moreover, I keep track of subsets of the state space at which agents are entrepreneurs like so: $\mathcal{A}_{\mathcal{E}} := \{\, a_i \in \mathcal{A}$: agent i is an entrepreneur$\} \subset \mathcal{A}$; $\mathcal{Z}_{\mathcal{E}} := \{\, z_j \in \mathcal{Z}$: agent j is an entrepreneur$\} \subset \mathcal{Z}$; $\mathcal{H}_{\mathcal{E}} := \{h_k \in \mathcal{H}$: agent k is an entrepreneur$\} \subset \mathcal{H}$.

Proof of Proposition 1**:** The object of interest is the fraction of the labor force with human capital $h = j \in h_{min} \ldots h_{max}$ that opts into entrepreneurship. First, we can break this fraction down to the sum of unconstrained ($\mathcal{E}^{uc}$) and capital-constrained ($\mathcal{E}^{c}$) entrepreneurs, as the two sets are disjoint,

$$\frac{\mathcal{P}[\mathcal{E} \cap h = k]}{\mathcal{P}[h = k]} = \frac{\mathcal{P}[(\mathcal{E}^{uc} \cup \mathcal{E}^{c}) \cap h = k]}{\mathcal{P}[h = k]} = \frac{\mathcal{P}[(\mathcal{E}^{uc} \cap h = k]}{\mathcal{P}[h = k]} + \frac{\mathcal{P}[(\mathcal{E}^{c} \cap h = k]}{\mathcal{P}[h = k]} \tag{A10}$$

Using the rules of conditional probability and the conditions for occupational choice,

$$\begin{aligned}
\frac{\mathcal{P}[\mathcal{E} \cap h = k]}{\mathcal{P}[h = k]} &= \frac{\mathcal{P}[(\mathcal{E}^{uc}|h = k]\,\mathcal{P}[h = k]}{\mathcal{P}[h = k]} + \frac{\mathcal{P}[(\mathcal{E}^{c} \cap h = k]\,\mathcal{P}[h = k]}{\mathcal{P}[h = k]} \\
&= \mathcal{P}[(\mathcal{E}^{uc} \,|\, h = k] + \mathcal{P}[(\mathcal{E}^{c} \,|\, h = k] \\
&= \mathcal{P}\left[z > \underline{z}^{uc}(h_k) \cap a_i > a^*(z, h_k) \,|\, h = k\right] + \mathcal{P}\left[z > \underline{z}^{c}(a_i, h_k) \cap a_i \leq a^*(z, h_k) \,|\, h = k\right] \\
&= \mathcal{P}\left[z > \underline{z}^{uc}(h_k) \,|\, h = k\right]\ \mathcal{P}[a_i > a^*(z, h_k) \,|\, z > \underline{z}^{uc}(h_k), h = j] \\
&\quad + \mathcal{P}\left[z > \underline{z}^{c}(a_i, h_k) \,|\, h = k\right]\ \mathcal{P}[a_i \leq a^*(z, h_k) \,|\, z > \underline{z}^{c}(a_i, h_k), h = k] \\
&= (\mathbf{1} - \boldsymbol{\Lambda_k})'(\mathbf{1} - \boldsymbol{G_z}[\underline{\boldsymbol{z}}^{\boldsymbol{uc}}(\boldsymbol{h_k})]) + \boldsymbol{\Lambda_k}'(\mathbf{1} - \boldsymbol{G_z}[\underline{\boldsymbol{z}}^{\boldsymbol{c}}(\boldsymbol{a_i}, \boldsymbol{h_k})])
\end{aligned} \tag{A11}$$

where the vector of $\mathbf{\Lambda}$'s and complementary CDF's are defined accordingly to express the previous equality as a non-negative linear combination ($0 \leq \Lambda_i \leq 1$) of non-negative functions ($0 \leq (1 - G_z(z_j)) \leq 1$) for every pair of (a_i, z_j). The cutoffs for the tail distributions are determined by the conditions given in Lemma 2. It is now simple to derive necessary and sufficient conditions for A11 to be strictly decreasing/increasing in the domain of human capital.

• ($\mathit{\Lambda} = \mathit{0}$ or $\lambda \to \infty$) This is the case when equation **??** fully characterizes the solution. The cutoff $\underline{z}^{uc}(h)$ is weakly increasing in human capital for $h \leq \hbar$, and for $\omega > \nu$, it is weakly decreasing for $h > \hbar$. If the marginal CDF of z is strictly increasing (no flat intervals), then the object of interest is strictly decreasing for $h \leq \hbar$ and strictly increasing for $h > \hbar$. Therefore, the necessary and sufficient condition is simply $\omega > \nu$.

• ($\mathit{\Lambda} = \mathit{1}$) This is the case when equation **??** fully applies and is analogous to the previous one with the exception that the parameter restriction is now given by $\omega > \nu + \alpha - \alpha\nu$. The necessary and sufficient condition in this case is simply $\omega > \nu + \alpha - \alpha\nu$.

• ($\mathit{0} \leq \mathit{\Lambda} \leq \mathit{1}$) The general case is analyzed using the continuity of A11. All functions are decreasing for $h \leq \hbar$, so we only need to analyze the region above the threshold. For a given level of financial frictions λ and thus a vector $\boldsymbol{\Lambda_k}$ for every h, there exists some $\widehat{\omega} \in [\nu, \nu + \alpha - \alpha\nu]$ such that both components of A11 are strictly increasing in h for all $h > \hbar$. Since any non-negative linear combination of non-negative decreasing (increasing) functions is itself decreasing (increasing), and since both components have the same minimum in $\mathcal{H}$, it follows that A11 is a U-shaped function attaining its global minimum at $\hbar$, if and only if $\omega > \widehat{\omega}$. ■

Proof of Proposition 2**:** The simplest way to prove each result is to bound each function (capital demand, labor demand, profits) from below by another function that differs only by a positive additive term or a positive multiplicative constant. If this function is strictly increasing w.r.t. human capital, so is the function of interest. To save space I will illustrate the proof for the case of profit functions, with the procedure for every other case being analogous.

Begin by observing that the indirect profit function of an entrepreneur, whether constrained or unconstrained, is always increasing and convex in z. We can thus directly apply Jensen's inequality, as z_t is the only random variable for fixed h_e and a_t. Using this fact together with the results of Lemma 2 we get,

$$\begin{aligned}
\mathbb{E}\left[\pi^{uc}(z) \,|\, z > \underline{z}^{uc}(h_e), h_e\right] &\geq \pi^{uc}\big(\,\mathbb{E}\left[z \,|\, z > \underline{z}^{uc}(h_e), h_e\right]\big) \\
&= \mathbb{E}\left[z \,|\, z > \underline{z}^{uc}(h_e), h_e\right]^{\frac{1}{\nu}} \tilde{h}^{\frac{1}{\nu}}\, \nu \left[(1-\nu)\left(\frac{\alpha}{r+\delta}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{1-\nu}{\nu}} \\
&> \left(\underline{z}^{uc}(h_e)\right)^{\frac{1}{\nu}}\ \tilde{h}^{\frac{1}{\nu}}\, \nu \left[(1-\nu)\left(\frac{\alpha}{r+\delta}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{1-\nu}{\nu}} \\
&= h_e w \quad (\forall h_e \in \mathcal{H}) \qquad \text{(A12)}
\end{aligned}$$

where the strict inequality hold for all non $\mathbb{P}$-null sets. Similarly, for constrained entrepreneurs for any given $a \in \mathcal{A}$,

$$\begin{aligned}
\mathbb{E}\left[\pi^c(z) \,|\, z > \underline{z}^c(h_e), h_e\right] &\geq \pi^c\big(\mathbb{E}\left[z \,|\, z > \underline{z}^c(h_e), h_e\right]\big) \\
&= \mathbb{E}\left[z \,|\, z > \underline{z}^c(h_e), h_e\right]^{\frac{1}{1-\upsilon}} (1-\upsilon)\left[\widetilde{h}\,(\lambda a)^{\alpha(1-\nu)}\left(\frac{\upsilon}{w}\right)^{\upsilon}\right]^{\frac{1}{1-\upsilon}} - (\lambda a)(r+\delta) \\
&> \left(\underline{z}^c(h_e)\right)^{\frac{1}{1-\upsilon}} (1-\upsilon)\left[\widetilde{h}\,(\lambda a)^{\alpha(1-\nu)}\left(\frac{\upsilon}{w}\right)^{\upsilon}\right]^{\frac{1}{1-\upsilon}} - (\lambda a)(r+\delta) \\
&= h_e w C_a - (\lambda a)(r+\delta) \quad (\forall h_e \in \mathcal{H}) \qquad \text{(A13)}
\end{aligned}$$

where the term $C_a > 0$ is defined accordingly and again the strict inequality hold for all non $\mathbb{P}$-null sets. Since both A12 and A13 are strictly increasing in h_e, so is any positive linear combination of them. ■

Note that throughout the following derivations it is allowed to switch from multiple to repeated integration by the Fubini-Tonelli theorem: given the σ-finiteness of measure spaces and the measurability of functions, the iterated integrals are well-defined with respect to the unique product measure and equal to the multiple integrals.

Proof of Lemma 3**:** Aggregate physical capital (K) and aggregate human capital (H) can be derived as follows:

$$\begin{aligned}
K_* &= \iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} k(h_e, a, z)\, dG(h,a,z) \\
&= \iint\limits_{\mathcal{H}_{\mathcal{E}}\times\mathcal{Z}_{\mathcal{E}}} (z\widetilde{h}_e)^{\frac{1}{\nu}} \left(\frac{\alpha(1-\nu)}{r+\delta}\right)^{\frac{1-(1-\alpha)(1-\nu)}{\nu}} \left(\frac{(1-\alpha)(1-\nu)}{w}\right)^{\frac{(1-\alpha)(1-\nu)}{\nu}} dG(h,z) \\
&= \iint\limits_{\mathcal{H}_{\mathcal{E}}\times\mathcal{Z}_{\mathcal{E}}} (z\widetilde{h}_e)^{\frac{1}{\nu}}\, dG(h,z)\;\; X^{\frac{1-(1-\alpha)(1-\nu)}{\nu}}\;\Psi^{\frac{(1-\alpha)(1-\nu)}{\nu}} \qquad \text{(A14)}
\end{aligned}$$

$$\begin{aligned}
H_* &= \iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} \ell(h_e, a, z)\, dG(h,a,z) \\
&= \iint\limits_{\mathcal{H}_{\mathcal{E}}\times\mathcal{Z}_{\mathcal{E}}} (z\widetilde{h}_e)^{\frac{1}{\nu}} \left(\frac{\alpha(1-\nu)}{r+\delta}\right)^{\frac{\alpha(1-\nu)}{\nu}} \left(\frac{(1-\alpha)(1-\nu)}{w}\right)^{\frac{1-\alpha(1-\nu)}{\nu}} dG(h,z) \\
&= \iint\limits_{\mathcal{H}_{\mathcal{E}}\times\mathcal{Z}_{\mathcal{E}}} (z\widetilde{h}_e)^{\frac{1}{\nu}}\, dG(h,z)\; X^{\frac{\alpha(1-\nu)}{\nu}}\,\Psi^{\frac{1-\alpha(1-\nu)}{\nu}} \qquad \text{(A15)}
\end{aligned}$$

where for notational simplicity I define $X \equiv \frac{\alpha(1-\nu)}{r+\delta}$ and $\Psi \equiv \frac{(1-\alpha)(1-\nu)}{w}$. Therefore, the average level of entrepreneurial profit is strictly increasing in their human capital.

Then, aggregate income (real GDP) can be expressed as:

$$
\begin{aligned}
Y_* &= \iiint\limits_{\mathcal{H}\times\mathcal{A}\times\mathcal{Z}} z\widetilde{h}\left(k(h_e,a,z)^{\alpha}\,\ell(h_e,a,z)^{1-\alpha}\right)^{1-\nu} dG(h,a,z) \\
&= \iint\limits_{\mathcal{H}\times\mathcal{Z}} (z\widetilde{h})^{\frac{1}{\nu}}\, X^{\frac{\alpha(1-\nu)}{\nu}}\, \Psi^{\frac{(1-\alpha)(1-\nu)}{\nu}}\, dG(h,z) \\
&= \left(\iint\limits_{\mathcal{H}\times\mathcal{Z}} (z\widetilde{h})^{\frac{1}{\nu}}\, dG(h,z,\mathcal{E})\right)^{\nu} \left(\iint\limits_{\mathcal{H}\times\mathcal{Z}} (z\widetilde{h}_e)^{\frac{1}{\nu}}\, dG(h,z,\mathcal{E})\right)^{1-\nu} X^{\frac{\alpha(1-\nu)}{\nu}}\, \Psi^{\frac{(1-\alpha)(1-\nu)}{\nu}} \\
&= \left(\int\limits_{\mathcal{H}_{\mathcal{E}}}\int\limits_{\mathcal{Z}_{\mathcal{E}}} (z\widetilde{h})^{\frac{1}{\nu}}\, dG(h,z,\mathcal{E})\right)^{\nu} \left(K_*^{\alpha} H_*^{1-\alpha}\right)^{1-\nu} \\
&= \left(\int\limits_{\mathcal{H}_{\mathcal{E}}}\int\limits_{\mathcal{Z}_{\mathcal{E}}} z^{\frac{1}{\nu}}\,\widetilde{h}^{\frac{1}{\nu}}\, g(z,h|\mathcal{E})\, g(\mathcal{E})\, dz\, dh\right)^{\nu} \left(K_*^{\alpha} H_*^{1-\alpha}\right)^{1-\nu} \\
&= \left(1 - G_{\mathcal{Z}_{\mathcal{E}}\mathcal{H}_{\mathcal{E}}}(z,h)\right)^{\nu} \left(\int\limits_{\mathcal{H}_{\mathcal{E}}}\int\limits_{\mathcal{Z}_{\mathcal{E}}} z^{\frac{1}{\nu}}\,\widetilde{h}^{\frac{1}{\nu}}\, g(z,h|\mathcal{E})\, dz\, dh\right)^{\nu} \left(K_*^{\alpha} H_*^{1-\alpha}\right)^{1-\nu} \\
&= \left(1 - G_{\mathcal{Z}_{\mathcal{E}}\mathcal{H}_{\mathcal{E}}}(z,h)\right)^{\nu} \left(\mathbb{E}\left[(z\widetilde{h})^{\frac{1}{\nu}}|\mathcal{E}\right]\right)^{\nu} \left(K_*^{\alpha} H_*^{1-\alpha}\right)^{1-\nu}
\end{aligned}
\tag{A16}
$$

■

The purpose of Lemma 3 is to establish an important theoretical result in this setting, and to highlight the necessity of certain boundary conditions so that the result actually holds. The proof is instructive and straightforward as in Gabaix et al. (2016).

Proof of Lemma 4: The domain of both $\mathcal{A}$ and $\mathcal{B}$ is the Hilbert space $\mathcal{L}^2(\Omega)$: the space of square-integrable continuous functions v equipped with the appropriate inner product. In our case, the inner product between two real continuous trivariate functions $v(\mathbf{x})$ and $w(\mathbf{x})$ is defined as $\langle v, w\rangle = \iiint\limits_{\Omega} v(\mathbf{x})\, w(\mathbf{x})\, d^3\mathbf{x}$. To prove the result we need to show that $\langle v, \mathcal{B}g\rangle = \langle \mathcal{A}v, g\rangle$, for all $v \in \mathcal{L}^2(\Omega)$, which implies that $\mathcal{B} = \mathcal{A}^*$. As before, note the key use of the Fubini-Tonelli theorem.

$$
\begin{aligned}
\langle v, \mathcal{B} g \rangle &= \iiint\limits_{\Omega} v \left[-\frac{\partial}{\partial a}(\widetilde{s}\, g) - \frac{\partial}{\partial z}(\widetilde{\mu}_z\, g) + \frac{1}{2}\frac{\partial^2}{\partial z^2}(\widetilde{\sigma}_z^2\, g) - \frac{\partial}{\partial h}(\widetilde{\mu}_h\, g) + \frac{1}{2}\frac{\partial^2}{\partial h^2}(\widetilde{\sigma}_h^2\, g) \right] d^3\mathbf{x} \\
&= \underbrace{\int\limits_{\mathcal{H}} \int\limits_{\mathcal{Z}} - \Big[v\,(\widetilde{s}\, g) \Big]_{\mathcal{A}} dz\, dh}_{\Delta 1} + \int\limits_{\mathcal{H}} \int\limits_{\mathcal{Z}} \int\limits_{\mathcal{A}} (v_a\, \widetilde{s}\, g)\, da\, dz\, dh \\
&+ \underbrace{\int\limits_{\mathcal{H}} \int\limits_{\mathcal{A}} \left[v \left(-\widetilde{\mu}_z\, g + \frac{1}{2}\frac{\partial}{\partial z}(\widetilde{\sigma}_z^2\, g) \right) \right]_{\mathcal{Z}} da\, dh}_{\Delta 2} + \int\limits_{\mathcal{H}} \int\limits_{\mathcal{A}} \int\limits_{\mathcal{Z}} v_z \left(\widetilde{\mu}_z\, g - \frac{1}{2}\frac{\partial}{\partial z}(\widetilde{\sigma}_z^2\, g) \right) dz\, da\, dh \\
&+ \underbrace{\int\limits_{\mathcal{A}} \int\limits_{\mathcal{Z}} \left[v \left(-\widetilde{\mu}_h\, g + \frac{1}{2}\frac{\partial}{\partial h}(\widetilde{\sigma}_h^2\, g) \right) \right]_{\mathcal{H}} dz\, da}_{\Delta 3} + \int\limits_{\mathcal{A}} \int\limits_{\mathcal{Z}} \int\limits_{\mathcal{H}} v_h \left(\widetilde{\mu}_h\, g - \frac{1}{2}\frac{\partial}{\partial h}(\widetilde{\sigma}_h^2\, g) \right) dh\, dz\, da \\
&= \iiint\limits_{\Omega} (v_a\, \widetilde{s})\, g + (v_z\, \widetilde{\mu}_z)\, g + (v_h\, \widetilde{\mu}_h)\, g\; d^3\mathbf{x} \\
&+ \underbrace{\int\limits_{\mathcal{H}} \int\limits_{\mathcal{A}} \left[-\frac{1}{2}\, \widetilde{\sigma}_z^2\, g\, v_z \right]_{\mathcal{Z}} da\, dh}_{E1} + \int\limits_{\mathcal{H}} \int\limits_{\mathcal{A}} \int\limits_{\mathcal{Z}} \frac{1}{2}\, \widetilde{\sigma}_z^2\, g\, v_{zz}\, dz\, da\, dh \\
&+ \underbrace{\int\limits_{\mathcal{A}} \int\limits_{\mathcal{Z}} \left[-\frac{1}{2}\, \widetilde{\sigma}_h^2\, g\, v_h \right]_{\mathcal{H}} dz\, da}_{E2} + \int\limits_{\mathcal{A}} \int\limits_{\mathcal{Z}} \int\limits_{\mathcal{H}} \frac{1}{2}\, \widetilde{\sigma}_h^2\, g\, v_{hh}\, dh\, dz\, da \\
&= \iiint\limits_{\Omega} (v_a\, \widetilde{s})\, g + (v_z\, \widetilde{\mu}_z)\, g + (v_h\, \widetilde{\mu}_h) g + \frac{1}{2}\, (v_{zz}\, \widetilde{\sigma}_z^2)\, g + \frac{1}{2}\, (v_{hh}\, \widetilde{\sigma}_h^2) g\; d^3\mathbf{x} \\
&= \langle \mathcal{A} v, g \rangle
\end{aligned}
$$

I use integration by parts to go from the first to the second equality, and once again to arrive to the third equality. The condition $(\Delta 1 + \Delta 2 + \Delta 3) = 0$ is satisfied because of the boundary condition (22) , which comes from integrating the KFE over Ω. To obtain the final result we need $E1 = 0$ and $E2 = 0$, which result in the general boundary conditions $v_z(h, a, \underline{z}) = v_z(h, a, \bar{z}) = 0$ and $v_h(\underline{h}, a, z) = v_h(\bar{h}, a, z) = 0$. When $v(.)$ is the value function $V(h, a, z)$, these boundary conditions are automatically satisfied because the reflecting barriers in $[\underline{z}, \bar{z}]$ and $[\underline{h}, \bar{h}]$ imply the SCBCs and NBCs (17) through (19). ■

Appendix B: Numerical Methods

To solve the model numerically I use an *implicit upwind finite difference* scheme similar to Achdou et al. (2021).[19] In this part of the appendix I summarize the most important features of the employed numerical methods and refer the reader to the cited manuscripts throughout the text for even more details.

B1. Construction of Grids

The value functions are approximated in the joint state space $(\mathcal{A} \times \mathcal{Z} \times \mathcal{H})$, discretized at $(I \times J \times K)$ grid points. I use a standard uniform grid for the dimension of entrepreneurial productivity. The process for z_t is approximated on the interval $[0.6696, 4.4260]$ at 40 equidistant grid points, where the reflecting barriers are chosen such that $G_z(\underline{z}) = 0.2000$ $G_z(\bar{z}) = 0.9991$, with $G_z(.)$ being the (stationary) lognormal distribution under the calibrated mean and variance. The process for h_t is necessarily approximated on a non-uniform grid in view of the log piecewise-linear mapping between education and human capital 31. The "full grid" would consist of 22 points, but in order to reduce the considerable computational cost of solving the model I choose an 11-point grid, effectively skipping every other year of education: $\widetilde{S} = 0, 2, 4, \ldots, 20$. The lower and upper bound of the grid are determined by the mapping itself: $\underline{h} = 1$ and $\bar{h} = 5.789$.

I also work with a non-equispaced grid of 550 points for the dimension of assets (a_t). It is quite known in the field of computational physics as well as in quantitative finance, e.g., (Tavella and Randall (2000), that it is useful to adapt the grid to the nature of the problem at hand in order to drastically improve the accuracy of finite difference calculations. Moreover, it is an established result that in model economies similar to ours the value function displays significant curvature in low-wealth regions, especially near the borrowing constraint, whereas it becomes approximately linear for larger values of wealth. To concentrate the mesh in areas where the value function is known to be more sensitive, I use an affine transformation to construct a grid that is around 20 times more dense (compared to a uniform grid) in its first third, around 10 times more dense towards the middle, and progressively more sparse for larger asset holdings. Finally, there are two actions I take to determine an appropriate upper bound for assets: the resulting grid covers the full support of the stationary wealth distribution, and the upper bound does not affect the saving decisions of the highest-type agents with non-zero measure. One of the lowest upper bounds that satisfies the above is $\bar{a} = 11,500$.

[19] As discussed in the text, a good list of references for the numerical methods I employ in this paper is, and especially the main text and numerical appendix of Achdou et al. (2021), which I follow closely. My computer program is also based on the very useful set of codes and notes of Benjamin Moll and his collaborators, which can be found here: `https://benjaminmoll.com/codes/`.

I denote by Δx^+ (Δx^-) the forward (backward) inter-grid distances for each variable $x = h, a, z$. To make notation easier to read I use the shorthand $V_{i,j,k} := V(a_i, z_j, h_k)$. At each point on the grid the first-order partial derivatives of $V_{i,j,k}$ are computed with either a *forward* or a *backward* finite difference approximation,

$$
\begin{aligned}
\partial_a^F V_{i,j,k} &:= \frac{V_{i+1,j,k} - V_{i,j,k}}{\Delta a^+} & \partial_a^B V_{i,j,k} &:= \frac{V_{i,j,k} - V_{i-1,j,k}}{\Delta a^-} \\
\partial_z^F V_{i,j,k} &:= \frac{V_{i,j+1,k} - V_{i,j,k}}{\Delta z^+} & \partial_z^B V_{i,j,k} &:= \frac{V_{i,j,k} - V_{i,j-1,k}}{\Delta z^-} \\
\partial_h^F V_{i,j,k} &:= \frac{V_{i,j,k+1} - V_{i,j,k}}{\Delta h^+} & \partial_h^B V_{i,j,k} &:= \frac{V_{i,j,k} - V_{i,j,k-1}}{\Delta h^-}
\end{aligned}
\tag{B1}
$$

and the second-order partial derivatives using a *central* difference approximation,

$$
\begin{aligned}
\partial_{zz} V_{i,j,k} &:= \frac{V_{i,j+1,k} - 2V_{i,j,k} + V_{i,j-1,k}}{(\widetilde{\Delta z})^2} \\
\partial_{hh} V_{i,j,k} &:= \frac{V_{i,j,k+1} - 2V_{i,j,k} + V_{i,j,k-1}}{(\widetilde{\Delta h})^2}
\end{aligned}
\tag{B2}
$$

where $(\widetilde{\Delta x})^2 := \frac{1}{2}(\Delta x^+ + \Delta x^-) \cdot (\Delta x^+ \cdot \Delta x^-)$. Note that in the simplest case of equispaced grids, $(\widetilde{\Delta x})^2$ reduces to simply $(\Delta x)^2$.

The backward and forward difference approximations for the agent's savings – the *optimal control drift* – are defined as

$$
\begin{aligned}
s_{i,j,k}^F &:= \mathcal{Y}_{i,j,k} + r a_i - (u')^{-1}(\partial_a^F V_{i,j,k}) \\
s_{i,j,k}^B &:= \mathcal{Y}_{i,j,k} + r a_i - (u')^{-1}(\partial_a^B V_{i,j,k})
\end{aligned}
\tag{B3}
$$

Analogously to the above, define the (discretized) forward and backward *Hamiltonians*,

$$
\begin{aligned}
H_{i,j,k}^F &:= u(c_{i,j,k}^F) + \partial_a^F V_{i,j,k}\ s_{i,j,k}^F \\
H_{i,j,k}^B &:= u(c_{i,j,k}^B) + \partial_a^B V_{i,j,k}\ s_{i,j,k}^B
\end{aligned}
\tag{B4}
$$

Non-convexities may result in value functions that are not strictly concave in the endogenous state variable. In fact, it gives rise to a *convex kink*, so problems may come up in the approximation of the optimal drift when we have both $s_{i,j,k}^F > 0$ and $s_{i,j,k}^B < 0$.

To deal with this case, a fast and reliable solution comes from the field of computational physics: a so-called *upwind scheme*. In such a scheme, the gist is to use a forward difference whenever the drift of each state variable is positive, and a backward difference whenever the drift is negative. As the authors explain, in our case the appropriate upwind scheme is the following approximation,

$$
\begin{aligned}
v'_{i,j} = \, & v'_{i,j,F} \left(\mathbb{1}_{\{s_{i,j,F}>0\}} \mathbb{1}_i^{unique} + \mathbb{1}_{\{H_{i,j,F} \geq H_{i,j,B}\}} \mathbb{1}_i^{both} \right) \\
& + v'_{i,j,B} \left(\mathbb{1}_{\{s_{i,j,B}<0\}} \mathbb{1}_i^{unique} + \mathbb{1}_{\{H_{i,j,F} \leq H_{i,j,B}\}} \mathbb{1}_i^{both} \right) \\
& + \bar{v}'_{i,j} \; \mathbb{1}_{\{s_{i,j,F} \leq 0 \leq s_{i,j,B}\}}
\end{aligned} \tag{B5}
$$

where the indicator $\mathbb{1}_i^{both}$ is defined for the "problematic case" in which both $s_{i,j,F} > 0$ and $s_{i,j,B} < 0$, and the indicator $\mathbb{1}_i^{unique}$ is defined for the "unproblematic" cases $s_{i,j,F} < 0$ and $s_{i,j,B} > 0$, as well as $s_{i,j,F} < 0$ and $s_{i,j,B} < 0$.

Putting it all together, the implicit upwind finite difference approximation reads,

$$
\begin{aligned}
\frac{V_{i,j,k}^{n+1} - V_{i,j,k}^{n}}{\Delta} + \rho V_{i,j,k}^{n+1} = \, & u\left(c_{i,j,k}^{n}\right) \\
& + \partial_a^F V_{i,j,k}^{n+1} \left(s_{i,j,k}^{F_n}\right)^+ + \partial_a^B V_{i,j,k}^{n+1} \left(s_{i,j,k}^{B_n}\right)^- \\
& + \partial_z^F V_{i,j,k}^{n+1} \left(\tilde{\mu}_{z_j}\right)^+ + \partial_z^B V_{i,j,k}^{n+1} \left(\tilde{\mu}_{z_j}\right)^- + \frac{1}{2} \tilde{\sigma}_{z_j}^2 \partial_{zz} V_{i,j,k}^{n+1} \\
& + \partial_h^F V_{i,j,k}^{n+1} \left(\tilde{\mu}_{h_k}\right)^+ + \partial_h^B V_{i,j,k}^{n+1} \left(\tilde{\mu}_{h_k}\right)^- + \frac{1}{2} \tilde{\sigma}_{h_k}^2 \partial_{hh} V_{i,j,k}^{n+1}
\end{aligned} \tag{B6}
$$

Superscripts denote the iteration counter, and the positive and negative part of the drift is expressed using the notation $f(x)^+ = \max\{f(x), 0\}$ and $f(x)^- = -\min\{f(x), 0\}$, respectively. Notice both the n superscript on the lhs and the $n+1$ superscripts on the rhs to see why the scheme is (semi)-implicit[20] keeping in mind that the HJB equation is solved backwards in time.

The result is an $(I \times J \times K)$-dimensional system of equations that can be expressed in matrix form and solved very efficiently using sparse matrix procedures. The first step is to substitute for the partial derivative approximations (B1) and (B2) into (B6), collect the terms with the same subscripts, and rearrange to get,

$$
\begin{aligned}
\frac{V_{i,j,k}^{n+1} - V_{i,j,k}^{n}}{\Delta} + \rho V_{i,j,k}^{n+1} = \, & u\left(c_{i,j,k}^{n}\right) \\
& + V_{i-1,j,k}^{n+1} \, \alpha_{i,j,k}^{n} + V_{i,j,k}^{n+1} \left(\beta_{i,j,k}^{n} + \chi_j^z + \chi_k^h\right) + V_{i+1,j,k}^{n+1} \, \gamma_{i,j,k}^{n} \\
& + V_{i,j-1,k}^{n+1} \, \phi_j^z + V_{i,j+1,k}^{n+1} \, \psi_j^z \\
& + V_{i,j,k-1}^{n+1} \, \phi_k^h + V_{i,j,k+1}^{n+1} \, \psi_k^h
\end{aligned} \tag{B7}
$$

The auxiliary variables $(\alpha^n, \beta^n, \gamma^n)$ are encoding information about agents' optimal savings, i.e., the drift of the *endogenous* state (in the absence of a diffusion term), at each iteration and each sets of grid points (i, j, k),

[20] In contrast to explicit finite difference schemes being computationally slow because of the Courant-Friedrichs-Lewy (CFL) condition, the employed implicit scheme achieves fast convergence because it allows for arbitrarily an large time step. Under value function iteration, the term $\frac{V_{i,j,k}^{n+1} - V_{i,j,k}^{n}}{\Delta} \rightarrow 0$ as $V_{i,j,k}^{n+1} \rightarrow V_{i,j,k}^{n}$.

$$\alpha_{i,j,k}^n := -\frac{\left(s_{i,j,k}^{B_n}\right)^-}{\Delta a^-},\ \beta_{i,j,k}^n := -\frac{\left(s_{i,j,k}^{F_n}\right)^+}{\Delta a^+} + \frac{\left(s_{i,j,k}^{B_n}\right)^-}{\Delta a^-},\ \gamma_{i,j,k}^n := \frac{\left(s_{i,j,k}^{F_n}\right)^+}{\Delta a^+} \tag{B8}$$

and the auxiliary variables $(\varphi^z, \varphi^h, \chi^z, \chi^h, \psi^z, \psi^h)$ are encoding information about the drift and diffusion of the two *exogenous* states at each grid point,

$$\begin{aligned}
\varphi_j^z &:= -\frac{(\tilde{\mu}_j^z)^-}{\Delta z^-} + \frac{\tilde{\sigma}_{z_j}}{2(\Delta z)^2} & \varphi_k^h &:= -\frac{(\tilde{\mu}_k^h)^-}{\Delta h^-} + \frac{\tilde{\sigma}_{h_k}}{2(\Delta h)^2} \\
\chi_j^z &:= \frac{(\tilde{\mu}_j^z)^-}{\Delta z^-} - \frac{(\tilde{\mu}_j^z)^+}{\Delta z^+} - \frac{\tilde{\sigma}_{z_j}}{(\Delta z)^2} & \chi_k^h &:= \frac{(\tilde{\mu}_j^h)^-}{\Delta h^-} - \frac{(\tilde{\mu}_j^h)^+}{\Delta h^+} - \frac{\tilde{\sigma}_{h_k}}{(\Delta h)^2} \\
\psi_j^z &:= \frac{(\tilde{\mu}_j^z)^+}{\Delta z^+} + \frac{\tilde{\sigma}_{z_j}}{2(\Delta z)^2} & \psi_k^h &:= \frac{(\tilde{\mu}_k^h)^+}{\Delta h^+} + \frac{\tilde{\sigma}_{h_k}}{2(\Delta h)^2}
\end{aligned} \tag{B9}$$

Using the above definitions, the discretized system can be represented in matrix form,

$$\frac{1}{\Delta}\left(\mathbf{v}^{n+1} - \mathbf{v}^n\right) + \rho\mathbf{v}^{n+1} = \mathbf{u}^n + \mathbf{A}^n\,\mathbf{v}^{n+1} \tag{B10}$$

where $\mathbf{v}^{n+1} = \mathrm{vec}(\mathbf{V}^{n+1}),\ \mathbf{v}^n = \mathrm{vec}(\mathbf{V}^n)$, and $u^n = \mathrm{vec}(\mathbf{u}^n)$ are the vectorized versions of the 3D value function and utility matrices; that is, they are column vectors of length $(I \times J \times K) \times 1$. The (very sparse) matrix $\mathbf{A}^n = \left(\widetilde{\mathbf{A}}^n + \widetilde{\mathbf{Z}} + \widetilde{\mathbf{H}}\right)$ is of size $(I \times J \times K) \times (I \times J \times K)$.

It is instructive to make the connection between formal theory and practice even more explicit. It is understood that the "HJB operator", $\mathcal{A}$, is the infinitesimal generator of the joint stochastic process (a, z, h). Such a differential operator can be thought of as the infinite-dimensional analogue of a continuous time transition matrix. The matrix $\mathbf{A}^n$ is exactly the discretized version of the infinitesimal generator $\mathcal{A}$, confined within the reflecting barriers of the process (a, z, h). As a result the elements of $\mathbf{A}^n$ check out the three properties that a proper *Markov transition rate matrix* (or *intensity matrix*) needs to satisfy. In particular, *i)* $0 \leq -q_{ii} \leq \infty$; *ii)* $0 \leq q_{ij},\ \forall\, i \neq j$; *iii)* $\sum_i q_{ij} = 0,\ \forall\, i$.

Finally, we can re-express the system as a *sparse system of linear equations* of the general form $\mathbf{Ax} = \mathbf{b}$. The matrix $\mathbf{B}^n$ is of dimension $(I \times J \times K) \times (I \times J \times K)$ and the vector $\mathbf{b}^n$ is of $(I \times J \times K) \times 1$, so the system can be solved for $\mathbf{v}^{n+1}$ in one step,

$$\mathbf{B}^n\mathbf{v}^{n+1} = \mathbf{b}^n \tag{B11}$$

where $\mathbf{B}^n := \left(\frac{1}{\Delta} + \rho\right)\mathbf{I} - \mathbf{A}^n$ and $\mathbf{b}^n := \mathbf{u}^n + \frac{1}{\Delta}\mathbf{v}^n$

B3. Computational Algorithm for the Stationary Equilibrium

I solve for the stationary equilibrium of the model using an extension of the methods described in Achdou et al. (2021). Among other things, the authors have developed a set of algorithms designed to obtain stationary equilibria in Aiyagari-type economies. The difference is that I have to iterate on two prices – the real interest rate (r) and wage rate (w), until both markets clear.

1) Guess the equilibrium interest rate r^l for $l = 0, 1, 2, \ldots$ on a reasonable interval $[r_{min}, r_{max}]$ and set an initial guess $r_{min} \leq r_0 \leq r_{max}$. Guess the equilibrium wage rate w^l for $l, m = 0, 1, 2, \ldots$ on a reasonable interval $[w_{min}, w_{max}]$ and set an initial guess $w_{min} \leq w_0 \leq w_{max}$.

2) Given factor prices r^l and w^l, solve the HJB equations and the KF equations using the implicit finite difference upwind scheme explained in the previous sections.

3) Compute the excess demand for capital/assets, $ED_a(r^l, w^l) = K(r^l, w^l) - S(r^l, w^l)$, the excess demand for labor/human capital, $ED_h(r^l, w^l) = L(r^l, w^l) - H(r^l, w^l)$, and check whether both markets clear. If not, update r^l and w^l accordingly. Repeat steps 2 & 3 until all markets clear under the conjectured prices.

Appendix C: Data Description; Statistical Procedures; Additional Tables

C1. Methodological Considerations for the SCF sample

As expected in such detailed surveys the data will involve some missing responses, either because the interviewee(s) did not provide some answers, or for confidentiality/non-disclosure reasons. The Federal Reserve imputes missing values using a statistical procedure called *multiple imputation*.[21] This procedure replaces each missing or deficient value with more acceptable values representing a realistic distribution of possibilities, with individual imputations generated by drawing repeatedly from an estimate of the conditional distribution of the data. Imputations are stored as five successive *implicates* of each data record. Thus, the total number of observations in each SCF wave is five times the actual number of responses. However, it is critical not to treat all five implicates as independent observations when conducting statistical analysis.

The SCF is based on a complicated structure and research design, which if ignored, it is guaranteed to yield errors when computing parameters and standard errors, and will seriously affect any hypothesis tests. The theory on proper inference in a multiple imputation setting is well-understood; it was developed by Donald B. Rubin in the 1970's and exemplified in his famous book Rubin (1987) (see also the survey of Little (1992), as well as Van Buuren (2018) for a good example of more recent work in the field). Below I summarize some of these issues and I explain how the current study addresses them.

1) When calculating sample statistics, estimating regressions, or carrying out any form of data analysis throughout the paper, the appropriate SCF survey weights are used for all observations. In other words, all reported statistics correspond to *weighted* sample estimates. This is essential in order to obtain accurate estimates while working with a truly nationally-representative sample. When I consider a number of robustness checks I also obtain unweighted estimates for many of the main results.

2) The SCF contains five implicates of every observation; each implicate may or many not have different values for any variable. Should a researcher decides to use all five implicates (which they should), this needs to be incorporated into calculations. To obtain valid point estimates one must compute the average point estimate across the five implicates. For instance, the correct sample mean/median/variance of a variable should equal the average of the means/medians/variances of the five implicates, or the correct $\hat{\beta}$ coefficient vector of an OLS regression should be computed as the average of the estimated $\hat{\beta}$ vectors after running the model five times using each of the implicates.

For the multivariate case, let $\widehat{\mathbf{Q}}$ be a $k \times 1$ vector of (properly weighted) estimates of the parameter vector $\mathbf{Q}$ that would have been obtained if no data were missing. Each

[21]The SCF research staff has developed the FRITZ (Federal Reserve Imputation Technique Zeta) model of multiple imputation, which incorporates insights from the Gibbs sampling algorithm together with statistical techniques developed in the image processing literature. See Kennickell (2017) for further details.

imputed data set $m = 1, \ldots, M$ admits an estimate of $\widehat{\mathbf{Q}}$, denoted $\widehat{\mathbf{Q}}_m$, along with a standard error $\sqrt{\mathbf{W}}_m$. The RII estimate of $\mathbf{Q}$ based on M imputed datasets is given by,

$$\overline{\mathbf{Q}} = \frac{1}{M} \sum_{m=1}^{M} \widehat{\mathbf{Q}}_m$$

3) In addition, when using all M complete-data versions of the potentially incomplete dataset and combine them into one pooled result, the uncertainty due to the missing data has to be taken into account. That is, we need to account for both *within-imputation* and *between-imputation* variability. Within variability is the sampling variance within each implicate; between variability is the variance contributed across implicates.

Specifically, within variability is given by the mean of the squared standard errors within the imputed data sets

$$\overline{\mathbf{W}} = \frac{1}{M-1} \sum_{m=1}^{M} \widehat{\mathbf{W}}_m$$

Between variability is given by the sample variance of the five parameters computed within each implicate

$$\mathbf{B} = \frac{1}{M-1} \sum_{m=1}^{M} \left(\widehat{\mathbf{Q}}_m - \overline{\mathbf{Q}}\right) \left(\widehat{\mathbf{Q}}_m - \overline{\mathbf{Q}}\right)'$$

Finally, the estimate of the total covariance matrix of the parameter vector is given by

$$\mathbf{V} = \overline{\mathbf{W}} + \left(1 + \frac{1}{M}\right) \mathbf{B}$$

4) Standard error calculations for the SCF call for special techniques. A very important consideration is how to compute correct standard errors in the presence of multiple imputation. If we naively treat each of the five implicates as an independent observation, the resulting standard errors will be substantially smaller and the corresponding hypothesis tests will overstate the statistical significance of the estimates. To give a rough idea about the empirical results of this paper, the calculated bootstrapped standard errors were typically larger by a factor of about $\sqrt{5}$ compared to traditional heteroskedasticity-robust standard errors.

5) The derivation of the degrees of freedom used in hypothesis testing and in generating confidence intervals is not straightforward. It is very likely that the SCF procedure under which the complete datasets are generated will affect the underlying degrees of freedom, thus ignoring it will result in artificially lower p-values and in more narrow confidence intervals. To deal with this issue, I follow Barnard and Rubin (1999) who propose a degrees-of-freedom correction to generate estimates that account for multiple imputation. It is worth noting that their proposed correction is conservative with respect to the standard method; it will always result in higher p-values and increase the width of confidence intervals.

Finally, all datasets come with a total of 999 replicates for each variable to simulate the complex sampling scheme of the SCF. The vector of replicate variables contains the number of times an observation was drawn for each replicate, and the bootstrap serves as a useful analog to the replicate process. For the first replicate, keep only the first implicate, and create a dataset with as many copies of each observation as are contained in the replicate variable. Then compute and store the desired parameter, using the newly expanded dataset. Repeat with another replicate variable.

C2. More Details on Identifying Entrepreneurs and Workers in the SCF

An essential difference between my analysis and Cagetti and De Nardi (2006) or De Nardi, Doctor and Krane (2007) is that they consider entrepreneurs in juxtaposition to the U.S. population, but instead I focus on households participating in the *labor force*. The need to appropriately compare economic outcomes between SEBO and non-SEBO requires excluding minors and people who are economically inactive, e.g., fully disabled persons, fully retired persons, volunteers, etc. I therefore designate a household to be part of the labor force if either the R or S/P satisfies the following requirements: **(i)** reports being economically active in some way, excluding volunteering; **(ii)** provides positive annual labor supply (number of weeks worked times hours worked per week); and **(iii)** declares non-zero pretax income. The subsequent analysis considers only households participating in the labor force, i.e., the sum of SEBO and non-SEBO.

Finally, I use the following procedure to decide whether the respondent or the spouse/partner should be considered the primary entrepreneur or worker. *1)* Assign SEBO status to either R or S/P according to which person has declared to engage in some form of self-employment. *2)* Assign SEBO status to the S/P if the R is marked as a non-participant in the operation of the business and at the same time the S/P participates in the operation of the business. Moreover, to wipe out the possibility of the R working at the business as a hobby, I assign SEBO status to the S/P is the R has documented no weeks of active work on their primary job in a normal year, whereas the S/P has a non-zero labor supply on their primary job as self-employed. *3)* If both persons are indeed self-employed in some way, then the S/P is assigned as SEBO if the person fulfills all of the above criteria and has more experience in the current business than R. Otherwise, R is given the assignment since it is more plausible for the more financially knowledgeable person with more experience to be a SEBO. Once this procedure is completed, the demographic and personal characteristics to be analyzed are assigned to be those of R or S/P.

Table C1: Weighted Probit Regressions; Probability of Being an Active Entrepreneur

	Dependent variable: SEBO $= 1$						
COVARIATES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Education	0.047***	0.028***	0.017***	-0.080^{***}	-0.036^{***}	-0.037^{***}	-0.039^{***}
	(0.0034)	(0.0035)	(0.0036)	(0.0122)	(0.0128)	(0.0128)	(0.0146)
Education^2 ($\times 100$)				0.347***	0.246***	0.248***	0.191***
				(0.0419)	(0.0436)	(0.0435)	(0.0501)
Potential experience					0.020***	0.033***	
					(0.0008)	(0.0026)	
Potential experience^2 ($\times 100$)						-0.0224^{***}	
						(0.0044)	
Self-Employed experience							0.073***
							(0.0016)
Worker experience							-0.002^{**}
							(0.0009)
ADDITIONAL CONTROLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Annual labor supply (log)			0.018	0.018	0.086***	0.076***	0.125***
			(0.0151)	(0.0151)	(0.0170)	(0.0171)	(0.0198)
Ever received inheritance (0/1)			0.245***	0.244***	0.156***	0.156***	0.125***
			(0.0199)	(0.0199)	(0.0200)	(0.0202)	(0.0230)
Expects to receive inheritance (0/1)			-0.010	-0.009	0.080***	0.078***	0.073***
			(0.0231)	(0.0229)	(0.0238)	(0.0240)	(0.0239)
Risk willingness ($1-4$)			0.133***	0.136***	0.168***	0.166***	0.174***
			(0.0094)	(0.0095)	(0.0096)	(0.0097)	(0.0101)
Health ($1-4$)		0.130***	0.119***	0.116***	0.153***	0.155***	0.132***
		(0.0177)	(0.0178)	(0.0178)	(0.0185)	(0.0185)	(0.0196)
Married (0/1)		0.499***	0.491***	0.490***	0.474***	0.457***	0.511***
		(0.0231)	(0.0236)	(0.0237)	(0.0239)	(0.0243)	(0.0258)
Male (0/1)		0.106***	0.088***	0.078***	0.083***	0.093***	-0.100^{***}
		(0.0220)	(0.0219)	(0.0220)	(0.0228)	(0.0232)	(0.0247)
Black (0/1)		-0.433^{***}	-0.393^{***}	-0.391^{***}	-0.346^{***}	-0.353^{***}	-0.302^{***}
		(0.0323)	(0.0329)	(0.0329)	(0.0336)	(0.0337)	(0.0368)
Hispanic (0/1)		-0.417^{***}	-0.366^{***}	-0.414^{***}	-0.291^{***}	-0.302^{***}	-0.254^{***}
		(0.0368)	(0.0358)	(0.0367)	(0.0384)	(0.0384)	(0.0429)
Other non-white (0/1)		-0.083^{***}	-0.047	-0.056^{*}	0.011	0.010	0.010
		(0.0310)	(0.0314)	(0.0313)	(0.0312)	(0.0314)	(0.0324)
Sector Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector $\times$ Occupation Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
McFadden's adjusted R^2	0.117	0.152	0.161	0.163	0.186	0.187	0.331
Observations	40, 413	40, 413	40, 413	40, 413	40, 413	40, 413	40, 413

Notes: All regressions are weighted by the appropriate SCF sampling weights. Estimated parameters are based on repeated imputation inference using all 5 SCF implicates. Bootstrapped standard errors (in parentheses) are calculated using all 999 SCF replicate draws and weights, and account for both imputation and sampling variability. All specifications include additional controls for *retirement*; *disability*; *employment variety*; *whether S/P is unemployed*; and a *constant term*, not reported here due to space limitations. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table C2: Weighted Least Squares Regressions; SEBO Hourly Business Income

	Dependent Variable: log hourly business income (> 0)						
COVARIATES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Education	0.079^{***}	0.076^{***}	0.075^{***}	-0.047	-0.040	-0.033	-0.047
	(0.0078)	(0.0078)	(0.0080)	(0.0332)	(0.0357)	(0.0319)	(0.0315)
Education^2 $(\times 100)$				0.417^{***}	0.393^{***}	0.346^{***}	0.368^{***}
				(0.1126)	(0.120)	(0.1081)	(0.1076)
Worker experience	0.007^{***}	0.008^{***}	0.009^{***}	0.009^{***}	0.006	0.003	0.002
	(0.0024)	(0.0024)	(0.0023)	(0.0023)	(0.0083)	(0.0075)	(0.0075)
Worker experience^2 $(\times 100)$					0.007	0.014	0.016
					(0.0222)	(0.020)	(0.020)
Self-empl. experience	0.018^{***}	0.017^{***}	0.019^{***}	0.018^{***}	0.046^{***}	0.039^{***}	0.028^{***}
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0057)	(0.0055)	(0.0055)
Self-empl. experience^2 $(\times 100)$					-0.070^{***}	-0.058^{***}	-0.043^{***}
					(0.0141)	(0.0138)	(0.0139)
Business employment size (log)						0.283^{***}	0.176^{***}
						(0.0185)	(0.0209)
Net business value (log)							0.167^{***}
							(0.0123)
ADDITIONAL CONTROLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ever received inheritance (0/1)			-0.081^{*}	-0.080^{*}	-0.089^{**}	-0.072^{*}	-0.045
			(0.0435)	(0.0434)	(0.0438)	(0.0432)	(0.0413)
Expects to receive inheritance (0/1)			-0.035	-0.032	-0.041	-0.008	0.014
			(0.0464)	(0.0470)	(0.0471)	(0.0501)	(0.0483)
Inherited any business (0/1)			0.221^{**}	0.235^{**}	0.229^{**}	0.181^{**}	0.051
			(0.0983)	(0.0989)	(0.0961)	(0.0868)	(0.0874)
Risk willingness $(1-4)$			0.073^{***}	0.079^{***}	0.080^{***}	0.062^{***}	0.027
			(0.0247)	(0.0247)	(0.0246)	(0.0237)	(0.0232)
Health $(1-4)$		0.140^{***}	0.137^{***}	0.135^{***}	0.122^{***}	0.107^{***}	0.097^{**}
		(0.0376)	(0.0377)	(0.0375)	(0.0381)	(0.0386)	(0.03888)
Male (0/1)		0.106^{**}	0.104^{**}	0.089^{**}	0.088^{*}	0.014	-0.065
		(0.0457)	(0.0458)	(0.0458)	(0.0465)	(0.0467)	(0.0473)
Black (0/1)		-0.240^{***}	-0.250^{**}	-0.244^{***}	-0.239^{***}	-0.276^{***}	-0.216^{**}
		(0.0886)	(0.0885)	(0.0885)	(0.0887)	(0.0898)	(0.090)
Hispanic (0/1)		-0.003	-0.007	-0.044	-0.055	-0.114	-0.122
		(0.0796)	(0.0795)	(0.0814)	(0.0809)	(0.0757)	(0.0786)
Other non-white (0/1)		0.078	0.070	0.057	0.047	0.047	0.046
		(0.0685)	(0.0687)	(0.0686)	(0.0683)	(0.0651)	(0.0650)
Sector Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector $\times$ Occupation Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Legal Entity Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,810	7,810	7,810	7,810	7,810	7,810	7,810
Adjusted R^2	0.149	0.154	0.157	0.159	0.165	0.215	0.249
RMSE	1.272	1.268	1.266	1.264	1.260	1.22	1.195

Notes: All regressions are weighted by the appropriate SCF sampling weights. Estimated parameters are based on repeated imputation inference using all 5 SCF implicates. Bootstrapped standard errors (in parentheses) are calculated using all 999 SCF replicate draws and weights, and account for both imputation and sampling variability. All specifications include additional controls for *firm ownership shares*; *number of firms owned and managed*; *marital status*; *retirement*; *disability*; *employment variety*; *whether S/P is unemployed*; and a *constant term*, not reported here due to space limitations. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table C1–Description of Variables Used in the Study

Variable	Measure of	SCF Code and Description
Age	Individual's Age	X14 (for R) and X19 (for S/P) in the Full Public Dataset.
Education	Years of Formal Education	X5901, X5905, X5931 (for R) and X6101, X6105, X6111 (for S/P) in the Full Public Dataset. Indicators for grades completed and highest degree earned are appropriately combined and converted into years of education, ranging from 0 to 21.
Potential Experience	Years of Potential Experience	Age - Education - 6, both defined as above.
Self-Employed Experience	Years of Experience Working Self-Employed	X4115, X4515, X4518, X4519, X4535, X4538, X4539 (for R) and X4715, X5115, X5115, X5118, X5119, X5135, X5138, X5139 (for S/P) in the Full Public Dataset. Variables are appropriately combined to calculate the years of past self-employed experience for non-SEBO and current self-employed experience for SEBO.
Worker Experience	Years of Experience Working as an Employee (not Self-Employed)	Potential Experience - Self-employed Experience, both defined as above.
Labor Income	Annual Labor Income	X5702 in the Full Public Dataset. *Q:* "In total, what was your annual income from wages and salaries, before deductions for taxes and anything else?"
Business Income	Annual Entrepreneurial Profit	X5704 + X5714 in the Full Public Dataset.*Q:* "In total, what was your net annual income from a sole proprietorship or a farm, before deductions for taxes and anything else?". *Q:* "In total, what was your annual income from other businesses or investments, net rent, trusts, or royalties, before deductions for taxes and anything else?"
Earned Income	Annual Earned Income	Labor Income + Business Income, both defined as above.
Active Business Value	Hired Capital by Entrepreneur	ACTBUS in the Summary Public Dataset. Net equity if (share of) businesses were sold today, plus loans from the household to business, minus loans from business to HH, plus value of personal assets used as collateral for business loans.
Number of Employees	Hired Labor by Entrepreneur	X3111 + X3211 in the Full Public Dataset. All paid workers in the businesses, both full-time and part-time, including the entrepreneur, members of his/her family, and anyone who is working without pay.
Net Worth	Total Wealth	NETWORTH in the Summary Public Dataset. The sum of total reported financial and non-financial wealth, minus total reported debt.
Net Worth excl. Business	Entrepreneurial Wealth Net of Business Interests	NETWORTH minus ACTBUS in the Summary Public Dataset. Net Worth minus the corresponding Active Business Value, both defined as above.
Business Share	Share of Firm owned by the household	X3128 and X3228 in the Full Public Dataset. *Q:* "What percentage of the business do you own?"
Labor Supply	Total hours of work in a typical year (hours per week × weeks per year)	X4110, X4111 (for R) and X4710, X4711 (for S/p) in the Full Public Dataset. *Q:* "How many hours do you work on your main job in a normal week?" *Q:* "Counting paid vacations as weeks of work, how many weeks do you work on your main job in a normal year?"
Employment Variety	Number of Different Past Employers	X4513 (for R) and X5113 (for S/p) in the Full Public Dataset. *Q:* "Including any self-employment and your current job, for how many different employers have you worked in full-time jobs lasting one year or more?"

Inherited Business (0/1)	Whether the entrepreneur has inherited any of his/her businesses or not	X3108=3 or X3108=4 or X3208=3 or X3208=4. *Q:* "How did you first acquire this business; was it bought or invested in, started by you, inherited, given to you, or some other way?"
Past Inheritance (0/1)	Whether the household has received any substantial inheritance in the past	X5801 in the Full Public Dataset. *Q:* "Have you ever received an inheritance, or been given substantial assets in a trust or in some other form? Please do not include inheritances from a deceased spouse."
Future Inheritance (0/1)	Whether the household expects to receive any substantial inheritance in the future	X5819 in the Full Public Dataset. *Q:* "Do you expect to receive a substantial inheritance or transfer of assets in the future?"
Turned Down (0/1)	R or S/P has been turned down for credit	X407=1 in the Full Public Dataset. *Q:* "In the past twelve months, has a particular lender or creditor turned down any request you made for credit, or not given you as much credit as you applied for?"
Not As Much (0/1)	R or S/P has not been given as much credit	X407=3 in the Full Public Dataset. *Q:* "In the past twelve months, has a particular lender or creditor turned down any request you made for credit, or not given you as much credit as you applied for?"
Fear Denial (0/1)	R or S/P feared of being turned down for credit	X409 in the Full Public Dataset. *Q:* "Was there any time in the past twelve months that you thought of applying for credit at a particular place, but changed your mind because you thought you might be turned down?"
Constrained (0/1)	R or S/P was likely financially constrained	Turned Down = 1 or Not as Much = 1, both defined as above.
Risk Willingness (1–4)	Willingness to take Financial Risk	X3014 (jointly for R and S/P) in the Full Public Dataset, recoded in inverse order. *Q:* "4) Take substantial financial risks expecting to earn substantial returns, 3) Take above average financial risks expecting to earn above average return, 2) Take average financial risks expecting to earn average returns, 1) Not willing to take any financial risks"
Health (1–4)		X6030 (for R) and X6124 (for S/P) in the Full Public Dataset, recoded in inverse order. *Q:* "Would you say your health in general is 4) excellent, 3) good, 2) fair, or 1) poor?"
Military (0/1)	Military Service Experience	X5906 (for R) and X6106 (for S/P) in the Full Public Dataset. *Q:* "(Have you/Has he or she) ever been in the military service?""
Retired (0/1)	Whether the individual has retired from a previous job	X4100=13 or X4100=50 if year $\leq$ 1992 and X6670=7 or X6671=7 … or X6677=7 if year$>$1992 (for R); X4700=13 or X4700=50 if year $\leq$ 1992 and X6678=7 or X6679=7 … or X6685=7 if year $>$1992 (for S/P) in the Full Public Dataset. Individual has reported being retired from a previous job.
Disabled (0/1)	Whether the individual has retired from a previous job	X4100=12 or X4100=52 if SCF year $\leq$ 1992 and X6670=6 or X6671=6 … or X6677=6 if SCF year $>$ 1992 (for R); X4700=12 or X4700=52 if SCF year $\leq$ 1992 and X6678=6 or X6678 =6 … or X6685=6 if SCF year $>$ 1992 (for S/P) in the Full Public Dataset. Individual has reported to have some disability.
Male (0/1)	Sex	X8021 (for R) and X103 (for S/P) in the Full Public Dataset. Individual has reported to be male or female.
Married (0/1)	Individual is either married or living with a partner	X8023 (for R) and X105 (for S/P) in the Full Public Dataset. *Q:* "Are you currently married or living with a partner, separated, divorced, widowed, or never been married?"
Unemployed Partner (0/1)	Individual is married/living with a partner who does not work	Married=1 and Labor Supply (of the other) = 0, both defined as above. *Q:* "Are you currently married or living with a partner, separated, divorced, widowed, or never been married?"

Race (1–4)	Individual's race	RACECL4 in the Summary Public Dataset. 1) White or Caucasian (includes Middle-Easter and Arab), 2) Black or African-American, 3) Hispanic or Latino, 4) Any other race.
No. of People	Number of People in the PEU	X101 in the Full Public Dataset.