The Macroeconomic Implications of Uncertainty and Learning for Entrepreneurship[*]

Han Gao[†]
University of New South Wales

Lichen Zhang[‡]
University of Hong Kong

Feb. 1, 2024

Abstract

We study the role of uncertainty and learning in selection into entrepreneurship at various life stages and its macroeconomic and policy implications. We develop a general equilibrium life cycle model of occupational choice under incomplete information about innate entrepreneurial ability. We discipline the learning process using novel subjective belief data on business forecasts of U.S. entrepreneurs. Our model aligns with key observed entrepreneurial life cycle outcomes, including the entry and exit patterns. It implies that the value of learning decreases with age, and switching to perfect information benefits those with the highest entrepreneurial ability the most. Quantitative experiments show that policies prioritizing the young by offering insurance for experimentation improve occupational sorting at earlier life stages, ultimately promoting aggregate entrepreneur share, output, and welfare.

Keywords: Entrepreneurship, Uncertainty, Imperfect Information, Learning, Life Cycle

JEL Classification: E13, E21, H25

[*]We are grateful to Anmol Bhandari, Mariacristina De Nardi, James Graham, Jonathan Heathcote, Hugo Hopenhayn, Zhen Huo, Loukas Karabarbounis, Mike Keane, Saigiri Kitao, Ellen McGrattan, Makoto Nakajima, Juanpa Nicolini, Sergio Ocampo Díaz, Fabrizio Perri, Todd Schoellman, Petr Sedláček, Venky Venkateswaran, Mike Waugh, Anson Zhou, Xiaodong Zhu, and participants at various seminars and conferences for their useful suggestions and comments. We thank the Australian Research Council (grants DP210103319 and CE170100005 for Gao) and the Hong Kong Research Grants Council (Early Career Scheme Grant 27503322 for Zhang) for the financial support. Earlier versions of this paper were circulated under the title "Personal Income Taxation and Entrepreneurship". All errors are our own. First version: August 2022.

[†]223 Anzac Pde Level 3, Kensington, NSW, 2033. Email: han.gao6@unsw.edu.au

[‡]901 K. K. Leung Building, University of Hong Kong, Pokfulam, Hong Kong. Email: lichenz@hku.hk

1 Introduction

A growing literature, dating back to the seminal work of Jovanovic (1982), emphasizes the importance of incorporating uncertainty and learning into models of entrepreneurship and firm dynamics, as they serve as underlying drivers of key macroeconomic variables such as growth, productivity, and welfare (Hopenhayn and Vereshchagina, 2009; David, Hopenhayn, and Venkateswaran, 2016; Kozlowski, Veldkamp, and Venkateswaran, 2020). In particular, uncertainty and learning play a crucial role in shaping the returns to entrepreneurship, as they generate value for experimentation, which directly influences the incentive for individuals to pursue entrepreneurial endeavors and create new businesses. Moreover, informational frictions can potentially interact with financial frictions, causing the relative importance of information uncertainty and learning in making entrepreneurial choices to differ at various life stages. This adds complexity to the policy design aimed at promoting entrepreneurship to reinvigorate the increasingly sclerotic U.S. business climate.

Despite numerous papers contributing to understanding the learning perspective of entrepreneurship,[1] very little is known about the quantitative importance of uncertainty arising from imperfect information and subsequent learning, as well as their interaction with other key factors, for selection into entrepreneurship at different stages of the life cycle, and, even more importantly, the corresponding macroeconomic and policy implications. This is mainly due to the lack of direct empirical evidence on entrepreneurial learning and the lack of a unified quantitative macroeconomic framework with sufficiently rich life cycle heterogeneity for comprehensive policy evaluation.

This paper aims to fill the gap in the literature by offering direct evidence on how entrepreneurs, under incomplete information, learn about their innate abilities and using it to discipline a quantitative Aiyagari-styled model with occupational choice and realistic life cycles, incorporating informational frictions, learning, and financial frictions. We use the model to quantify the value of learning and the cost of uncertainty arising from imperfect information, which informs the policy design for reviving entrepreneurship.

We show that incorporating life cycle learning dynamics under imperfect information into the model of entrepreneurial choice with incomplete markets changes people's incentives to become entrepreneurs at various life stages. Information uncertainty and learning provide individuals with incentive to select into entrepreneurship at a younger age to uncover their talents, which, interact with asset accumulation and financial frictions, help rationalize observed entry and exit dynamics for entrepreneurs over the life cycle and

[1] See, for example, Kerr, Nanda, and Rhodes-Kropf (2014), Manso (2016), and Hincapié (2020).

lead to distinct policy implications and welfare consequences. Policies targeting young individuals, by offering them insurance for experimentation with entrepreneurship, promote early resolution of uncertainty regarding their innate entrepreneurial abilities. This results in improved occupational sorting and ultimately increases the overall share of entrepreneurs, output, and welfare. Abstracting from informational frictions and learning significantly alters quantitative outcomes at both macro and micro levels.

We develop a general equilibrium heterogeneous agent life cycle overlapping generations model with occupational choice under imperfect information. At the end of each period, households choose for next period between being a worker and an entrepreneur based on heterogeneous characteristics, including occupation-specific abilities, beliefs, assets, and age. Workers earn wage income, while entrepreneurs earn business income in terms of profits from running their own businesses, subject to a collateral constraint on their own private savings. Both wage and business income are subject to a non-linear personal income tax schedule, *à* la Benabou (2002) and Heathcote, Storesletten, and Violante (2017). We also take other important determinants of occupational choice into consideration, including bequests and non-pecuniary motives.

The novel model block is uncertainty arising from imperfect information and learning about innate entrepreneurial ability, which is structured as follows. Agents are endowed with an unobserved permanent innate entrepreneurial ability upon entry into the labor market, which can be gradually learnt subject to noise only when working as entrepreneurs. The longer they remain entrepreneurs, the more accurately they discern their innate productivity.

We let the data speak to the relative importance of each factor as a determinant of entrepreneurial choice by exploiting the advantage of the method of simulated moments (MSM), which integrates information from various data sources, each of which is informative about certain aspects of U.S. entrepreneurs. We define entrepreneurs as self-employed pass-through business owners in our micro data. Drawing on the subjective belief data from the Panel Study of Entrepreneurial Dynamics (PSED), we find that: (i) entrepreneurs face significant uncertainty upon entry, resulting in large errors in forecasting their business performance; and (ii) they use their most recent information on sales to revise their future performance forecasts. To address the PSED's short panel limitations, we supplement it with the Panel Study of Income Dynamics (PSID), which allows us to document the key life cycle outcomes of entrepreneurs. In particular, we find the age profile of entrepreneurial entry is hump-shaped, with the entry rate peaking at middle age and the exit rate declining during the working age, and these patterns are also observed in richer administrative data. We use moments on the business forecasting updating process from the PSED to discipline the entrepreneurial productivity learning and further exploit the survey questions on personality traits (e.g., love of business) for estimating parameters that govern the non-pecuniary motives for being an entrepreneur.

While not directly targeted by the parameterization, our model is successful in matching salient features from both micro and macro data. In particular, it well replicates the hump-shaped entry pattern because in a finite life cycle, agents are incentivized to experiment with entrepreneurship early to reduce uncertainty about their future career paths. However, young individuals typically lack substantial wealth, making them more likely to be constrained from expanding there businesses. This, in turn, delays their entry into entrepreneurship. These opposing forces determine the timing of entrepreneurial entry, which contributes to generating realistic life cycle dynamics of entrepreneurship. The declining exit rate during the working age is mainly driven by the uncertainty and learning channel. The intuition behind is that only those who learn they have high innate entrepreneurial productivity stay in the market.

Armed with a well-fitted and validated model, we use it to deliver several key implications that will inform policy design. First, we find that the value of learning for an individual household, on average, decreases with age. Second, we quantify the cost of uncertainty arising from incomplete information by contrasting our benchmark model with a case of perfect information where agents perfectly know about their innate entrepreneurial ability upon entering the labor market. We show that the cost of informational frictions is higher for agents with higher innate ability, as switching to the case of perfect information benefits them more in terms of the share of lifetime spent as an entrepreneur and lifetime income. Third, informational frictions amplify financial frictions for the young. With perfect information, high-entrepreneurial-ability individuals enter the market immediately once the collateral constraint is eased. However, when entrepreneurial ability is unknown, individuals opt to accumulate extra wealth before entering, in case of a low realization of productivity shocks.

In light of the above results, resolving information uncertainty early enables individuals to discover their comparative advantage in entrepreneurial activities at a younger age. This disproportionately benefits those who are innately more productive, leading to improved occupational sorting and ultimately having a impact on aggregates. Therefore, policies aimed at promoting entrepreneurship should target the young, who are constrained by informational and financial frictions, by providing them with insurance for experimentation in their early career paths. To illustrate the idea, we consider two types of policies that have been widely proposed or even implemented in practice.

The first policy experiment is to directly subsidize entrepreneurs based on age. To make the experiments comparable, we fix the size of total subsidies across all cases. We find that subsidizing younger entrepreneurs effectively increases the entrepreneur population. Moreover, subsidizing all operating entrepreneurs generates a larger response relative to only subsidizing entrants. Our findings suggest that policies of direct subsidies aimed at boosting entrepreneurship should provide insurance for sufficiently long periods and target younger individuals. One example could be subsidizing young entrepreneurs for a few periods following their entry until uncertainty is largely resolved.

In our second policy experiment, we evaluate the impact of personal income tax progressivity on entrepreneurship by reverting the current U.S. progressive personal income tax scheme to a revenue-neutral flat tax in the entrepreneurial sector. This exercise is motivated by the observation that an age-dependent tax, as proposed by Weinzierl (2011) and Karabarbounis (2016), is challenging to implement in real-world practice, and a progressive tax scheme effectively mimics an age-dependent tax by placing a lower tax burden and offering higher insurance value to the young, who typically earn less due to high uncertainty and limited wealth.

Our findings show that the current U.S. progressive tax scheme is superior to the counterfactual flat business income tax reform in terms of promoting entrepreneurship. More specifically, the revenue-maximizing flat rate of around 20% achieves roughly the same tax revenue as the current U.S. progressive tax scheme. Even though switching to the revenue-neutral flat tax from the benchmark progressive tax reduces the average marginal tax rate faced by entrepreneurs from 26.0% to 24.1%, aggregate entrepreneur share decreases from 9.0% to 6.0%.

The flat tax reform not only discourages entrepreneurship at younger ages, but decreases entrepreneurship at older ages even more. The reasons are that a high flat tax rate discourages the young from entering entrepreneurship to discover their entrepreneurial aptitudes, and consequently, also discourages them from becoming entrepreneurs when they are older since the value of learning diminishes with age.

An even sharper message we want to convey is that it is people with the highest innate entrepreneurial productivity—who are more likely to establish high-growth startups ("gazelles")—that lose the most from the counterfactual flat tax reform. As highlighted in Decker, Haltiwanger, Jarmin, and Miranda (2016) and Sterk, Sedláček, and Pugsley (2021), these "gazelles" play a pivotal role in shaping aggregate firm growth and explaining overall business dynamism trends in the US. Due to the life cycle learning dynamics emphasized in our mechanism, without experimentation to learn about their innate entrepreneurial ability, the owners of those potential "gazelles" may never show up. This is, again, in contrast to a more conventional view that high-productivity entrepreneurs would benefit more from a flat tax compared to a progressive one. Consequently, the flat tax reform leads to a 1.6% drop in aggregate output, primarily due to reduced entrepreneurial production stemming from missing "gazelles," and a 2.0% decline in consumption-equivalent welfare. Abstracting the model from information uncertainty yields significantly different implications, as the flat tax reform would lead to a redistribution of lifetime income gains to agents with the highest entrepreneurial productivity.

Related Literature This paper contributes to several strands of literature. The first relates to the determinants of entrepreneurship and returns to self-employment, highlighting various aspects including wealth accumulation and financial frictions (Quadrini, 1999, 2000; Gentry and Hubbard, 2004; Cagetti and De Nardi, 2006; Buera, Kaboski, and Shin, 2011), non-pecuniary motives (Hamilton, 2000; Hurst and Pugsley, 2017), sweat equity (Bhandari and McGrattan, 2021), the risky nature of entrepreneurial activities (Hopenhayn and Vereshchagina, 2009; Boar and Knowles, 2020; Robinson, 2023).

In terms of entrepreneurship determinants, our paper is most related to the one focusing on learning and experimentation (Kerr, Nanda, and Rhodes-Kropf, 2014; Manso, 2016; Dillon and Stanton, 2018; Humphries, 2019; Hincapié, 2020; Hamilton, Hincapié, and Salari, 2022). We distinguish our work from those studies in three key aspects. First, instead of inferring uncertainty and learning process from indirect moments such as earnings and growth patterns—which are highly likely to attribute other unobserved heterogeneity to them—we use novel survey data on individual-level subjective forecasts of businesses as direct measures to discipline the learning process, and further show that our model effectively captures the earnings, entry, and exit dynamics of entrepreneurs over the life cycle.[2] This enhances the credibility and robustness of the core element of our quantitative model.

Second, we study the role of informational frictions in the presence of other key factors, particularly, asset accumulation subject to financial frictions, in a general-equilibrium setup with incomplete markets. We show that the interaction between the two frictions is essential for rationalizing the transitions in and out of entrepreneurship over the life cycle and and has significant implications for policy impacts. The advantage that our approach provides a unified framework encompassing all the key features is crucial. For example, early works came to mixed or even completely opposite conclusions regarding the impact of personal income tax progressivity on entrepreneurship due to different

[2] Using U.S. administrative data,Bhandari, Kass, May, McGrattan, and Schulz (2023) suggest that incorporating learning is essential to match young entrepreneurs' observed income growth profiles and switching behavior.

model assumptions and potential omissions of factors influencing entrepreneurial risk.[3] In contrast, the lack of a saving mechanism and general equilibrium features in aforementioned studies limits their capacity to evaluate the impacts of large-scale comprehensive policies on entrepreneurship, such as personal income tax.

Third, we quantify the importance of information uncertainty and learning for selection into entrepreneurship at various life cycle stages. This analysis provides a clearer rationale for encouraging experimentation and earlier resolution of uncertainty, identifies the types of policies that can promote entrepreneurship in practice, and assesses their aggregate and distributional impacts.

Our framework incorporating a realistic life cycle learning dynamics with imperfect information provides new angles for large-scale comprehensive policy evaluation of impacts on entrepreneurship. Previous macro and public finance literature has focused on various aspects of taxation and entrepreneurship such as top marginal tax rates (Imrohoroğlu, Kumru, and Nakornthab, 2018; Brüggemann, 2021), the role of business owner time in evaluating the impact of business income tax reform (Bhandari and McGrattan, 2021), wealth and capital income tax (Kitao, 2008; Boar and Knowles, 2020; Boar and Midrigan, 2023; Guvenen, Kambourov, Kuruscu, and Ocampo, 2023a; Guvenen, Kambourov, Kuruscu, Ocampo, and Chen, 2023b), the interaction with corporate tax (Cullen and Gordon, 2002; Sedláček and Sterk, 2019; Dyrda and Pugsley, 2020), and the redistribution and general equilibrium effects of personal income tax reform for both workers and business owners (Meh, 2005; Boháček and Zubrický, 2012).

In the classic heterogeneous agent occupational choice models with infinite horizon and purely exogenous entrepreneurial productivity processes (e.g., Evans and Jovanovic (1989), Quadrini (2000), Cagetti and De Nardi (2006), Buera, Kaboski, and Shin (2011)), the age of becoming an entrepreneur does not matter. Hence, policy considerations place more emphasis on high-income incumbent entrepreneurs. However, our model, which incorporates life cycle learning dynamics, emphasizes the importance of entrepreneurial experimentation for young individuals. We demonstrate that without young, talented agents becoming entrepreneurs first, older, large, successful business owners may never emerge. This key insight informs our primary policy implication regarding the impact of personal income tax on entrepreneurship: progressive taxation favors overall entrepreneurship, especially the most productive ones, by providing young individuals with insurance for experimentation, compared to a revenue-neutral flat tax. This chal-

[3] For instance, Wen and Gordon (2014) assume a log-normal income risk, leading to an increase in entrepreneur share with tax progressivity, while Fossen (2009) finds the opposite under the assumption of normal income risk.

lenges the conventional wisdom that progressive taxation discourages entrepreneurship by raising the expected marginal tax rate on high-income successful entrepreneurs.

Finally, our work complements the growing literature on the implications of information uncertainty and learning on firm dynamics, beginning with the seminal work of Jovanovic (1982). This framework has recently been extended to quantitative models of heterogeneous firms to examine the impact of imperfect information on misallocation and aggregate productivity (e.g., David, Hopenhayn, and Venkateswaran (2016), Senga (2018), Kozlowski, Veldkamp, and Venkateswaran (2020), and Chen, Senga, Sun, and Zhang (2023)), growth (Arkolakis, Papageorgiou, and Timoshenko, 2018), and the interaction between informational frictions and financial frictions (Kochen, 2022). As a complement to this literature highlighting the importance of incorporating learning for a firm's life cycle and industrial dynamics, we focus on how learning affects individual career choices and lifetime outcomes, as well as the corresponding policy implications.

As a cautious remark, we attribute the observed entry and exit dynamics of entrepreneurship over the life cycle to selection through learning about innate ability while abstracting from an important channel where entrepreneurs accumulate human capital to increase productivity, as emphasized in Bhandari and McGrattan (2021). Since these two channels have similar model implications in many aspects, previous attempts to distinguish them often rely on ad hoc assumptions or weak identification based on observational differences in limited attributes, such as in Nagypál (2007). Our decision to focus on information uncertainty and learning stems from the direct empirical evidence we find regarding the substantial uncertainty and belief updating process of business owners. This evidence helps to discipline our quantitative model. We thus view our choice as a natural starting point and acknowledge the potential for future extensions to disentangle these two channels with richer data.

2 Model

In this section, we develop a tractable quantitative general equilibrium model with rich heterogeneity that determines entrepreneurial choice decisions at different stages over the life cycle. The decision unit in the model represents the household head. We focus on a steady-state equilibrium, abstracting from time subscripts.

2.1 Demographics and Environment

Time is discrete, and there are three types of agents: the government, a representative corporate firm, and a continuum of heterogeneous households, populated by J overlapping generations. A model period corresponds to a year. In each period, a continuum of new households is born. The mass of cohorts grows at a rate of g_n. Each individual may die with an age-dependent probability ψ_j, with the conditional survival probability from age j to age $j+1$ denoted as $(1-\psi_j)$. The survival probability in the final period of the life cycle is $1-\psi_J = 0$. There are no annuity markets, and households who die derive warm-glow utility through leaving assets for future generations. Consequently, the economy features both accidental and voluntary bequests.

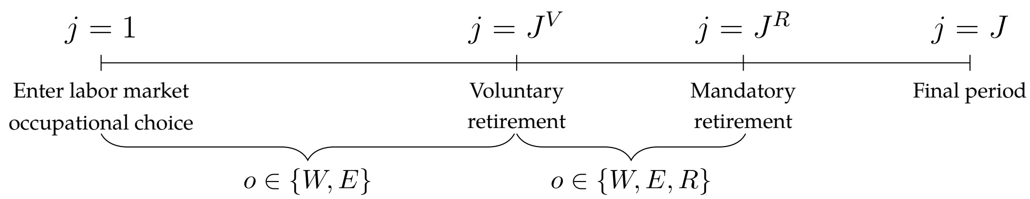

Figure 1: Life Cycle of the Model

Model age is indexed by $j \in \{1, ..., J\}$. We assume individuals younger than 21 are inactive. Model period 1 corresponds to age 21, when an individual enters the labor market. The youngest age to claim retirement and receive social security benefits is J^V, and the mandatory retirement age is J^R—after which all individuals must retire regardless of their current employment status. Hereafter, we refer to periods before J^V as normal working ages and periods between J^V and J^R as voluntary retirement ages.[4] For simplicity, our model abstracts from health shocks and medical expenditure shocks at older ages.

The time endowment for each individual household is one for each period. Before agents retire, they make occupational choices $o \in \{W, E\}$ at each period: to work as employees (W) or to own and operate a private business (E). The non-employed are interpreted as workers working zero hours in our model. At voluntary retirement ages, households choose $o \in \{W, E, R\}$ at each period, where R denotes retirees, which is an absorbing state.

We assume that all individual households enter the economy with zero initial assets (i.e., $a_0 = 0$). They receive a certain amount of bequest with some probability over the

[4]We include both voluntary and mandatory retirement options in our model to better align the age-profile of wealth and the exit rate of entrepreneurs around retirement age, as observed in the data. See Figure A11 and A12 in Appendix B of the Online Appendix for more details.

life cycle in an unanticipated way.[5] To account for the fact that large portion of lifetime income variation is determined before entering the labor market, as documented by Keane and Wolpin (1997), we assume households are ex ante endowed with a permanent love of business (LoB) characteristic x_e, a permanent worker skill type χ_w, and innate entrepreneurial ability μ, all of which are heterogeneous across households. While x_e and χ_w are perfectly known to households, μ is not observed. More specifically, when first entering the labor market, a household draws her own innate entrepreneurial productivity, μ, from a normal distribution $\mu \sim N\left(\mu_e, \nu_e^2\right)$, which is unobserved to her. Instead, a household forms a belief about the distribution of her innate ability and can only gradually learn about it through being an entrepreneur. We assume there is no learning on the workers' side.[6]

Households of age j make occupational choices based on idiosyncratic characteristics summarized by states $\mathbf{x}_j = (x_e,\ \chi_w,\ a_j,\ \epsilon_{w,j},\ \tilde{\mu}_{e,j},\ \tilde{\nu}_{e,j},\ \epsilon_{e,j})$, where x_e and χ_w are individual-fixed characteristics that have been introduced, a represents assets, ϵ_w is the idiosyncratic wage income shock in paid employment, and $(\ \tilde{\mu}_e,\ \tilde{\nu}_e,\ \epsilon_e)$ are age-dependent states governing the Bayesian learning process on a household's innate entrepreneurial productivity μ. The terms $\tilde{\mu}_e$ and $\tilde{\nu}_e$ denote the mean and standard error of the posterior beliefs regarding μ, and ϵ_e denotes a signal about μ. These learning process components will be discussed in full detail in the next section.

Before describing individual households' problems in detail, we outline the precise timing of the model within a period, as summarized in Figure 2. We assume that the occupation in period j is made at the end of period $j-1$. After choosing their occupation, those who decide to be entrepreneurs observe a signal about their innate entrepreneurial productivity, ϵ_e, at the beginning of period j. Given ϵ_e, the entrepreneur chooses capital and labor for production and updates their belief about their innate entrepreneurial productivity. After receiving incomes subject to taxes, they make decisions on consumption and saving. Finally, the entrepreneur decides on their occupation for the next period, and time evolves to time $j+1$. Conditional on being a worker, the individual receives wage income and makes consumption/saving decisions, as well as the occupational choice

[5]This simplifies our computation because the random draw does not enter individuals' expected value function.

[6]We admit that endogenous learning about ability occurs in all occupations, not just entrepreneurship, as in studies exploring dynamics of occupation choices, job transitions, and skill mismatch (Guvenen, 2007; Guvenen, Kuruscu, Tanaka, and Wiczer, 2020; Papageorgiou, 2014). We focus on learning exclusively within entrepreneurship to highlight that entrepreneurial activities are generally riskier and more uncertain compared to regular jobs (Hopenhayn and Vereshchagina, 2009; Boar, Gorea, and Midrigan, 2022; DeBacker, Panousi, and Ramnath, 2023), and workers can rapidly learn about their innate abilities (Guvenen et al., 2020). In particular, DeBacker, Panousi, and Ramnath (2023) show that business income is much riskier than labor income using a large panel of US income tax returns.

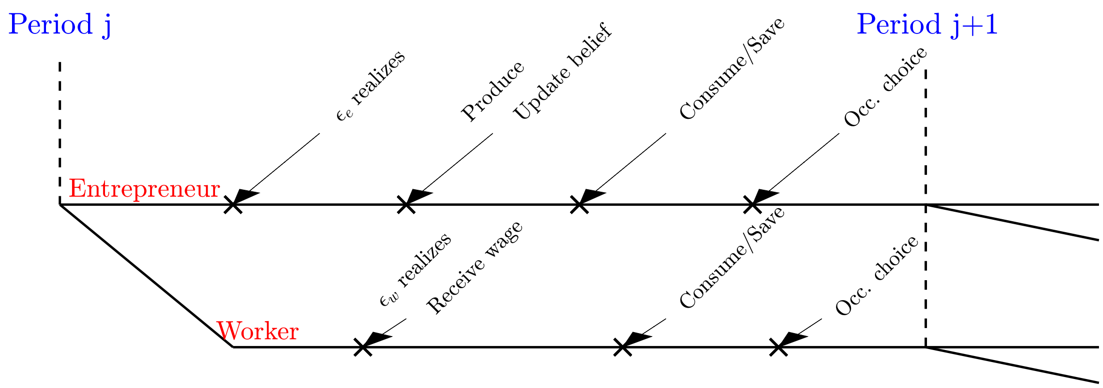

Figure 2: Timeline within One Model Period in Normal Working Ages

decision after the realization of idiosyncratic labor productivity shocks ϵ_w.

2.2 Preferences

Individual households maximize the expected utility function over sequences of consumption and leisure $\{c_j, l_j\}_{j=1}^{J}$:

$$\mathbb{E}\{\sum_{j=1}^{J} \beta^{j-1}[\prod_{i=0}^{j-1}(1-\psi_i)][(1-\psi_j)u\,(c_j, l_j) + \psi_j \mathcal{V}(a_j)]\},$$

where β is the discount factor, ψ is the age-dependent mortality shock, and $\mathcal{V}(a)$ denotes the value of the bequest. The expectation is taken with respect to the stochastic processes governing idiosyncratic labor productivity and learning about innate entrepreneurial productivity. Individual households' flow utility function is given as

$$u(c_j, l_j; x_e) = \frac{(c_j^{\gamma} l_j^{1-\gamma})^{1-\zeta}}{1-\zeta}, \quad \gamma \in (0,1),\ \zeta > 0$$

where γ is a utility weight of consumption and ζ determines individuals' risk aversion.

Both workers and entrepreneurs are endowed with one unit of productive time in every period. An individual household's leisure l_j is determined by

$$l_j = 1 - (\phi_{w,0} + h_j)\mathbb{I}_{\{h_j>0\}} - g(x_e)\mathbb{I}_{\{o_j=E\}}. \tag{1}$$

Workers of age $j < J^R$ split their time endowment between work h_j and leisure l_j.

If an individual chooses non-employment ($h = 0$), their leisure l is equal to one. If the individual works for positive hours (i.e., $h > 0$), besides the disutility from working hours, they suffer an extra fixed utility cost $\phi_{w,0}$.[7] Entrepreneurs ($o = E$) pay a fixed utility cost $g(x_e)$, denominated in productive time units. This cost is a function of the LoB state x_e, reflecting the non-pecuniary motive for entrepreneurship. Specifically, individuals with a higher x_e (e.g., those who enjoy being their own bosses) face a lower level of disutility as entrepreneurs. Hurst and Pugsley (2011, 2017) highlight the importance of non-pecuniary motives in entrepreneurial choice decisions. Hamilton, Papageorge, and Pande (2019) and Jones and Pratap (2020) demonstrate that non-pecuniary motives help explain the presence of low productivity and low-income entrepreneurs.

2.3 Entrepreneurial Productivity Learning Process

Upon entering the labor market, each individual household draws an innate entrepreneurial productivity μ from a normal distribution $\mu \sim \mathcal{N}(\mu_e, \nu_e^2)$ that is permanently associated with her. Since households do not have any information about their innate entrepreneurial abilities, their initial belief regarding μ follows the population distribution of entrepreneurial innate ability $\mathcal{N}(\mu_e, \nu_e^2)$.[8] Therefore, the initial uncertainty arises due to incomplete information regarding innate entrepreneurial ability.

Individual households learn about their innate entrepreneurial ability μ only after actively working as entrepreneurs. In every period as an entrepreneur, households receive an $\epsilon_{e,n}$ shock, which acts as a signal of their innate entrepreneurial productivity. Note that n captures the number of periods during which an individual household has worked as an entrepreneur. The signal consists of two parts: the innate entrepreneurial productivity μ and a transitory component ε that is independently and identically distributed across individual states and time, i.e., $\epsilon_{e,n} = \mu + \varepsilon_n, \;\; \varepsilon_n \sim i.i.d.\; \mathcal{N}(0, \sigma_e^2)$.

Note that both the innate entrepreneurial productivity and the transitory shocks are normally distributed, so the signal (or realized entrepreneurial productivity) $\epsilon_{e,n}$ is also normally distributed. Since both the priors and the signals are normally distributed, the distribution of the posterior beliefs after observing any number of signals will also be normally distributed. Hence, the distribution of the posterior beliefs about μ after observing the n-th signal can be completely described by its mean $\tilde{\mu}_{e,n}$ and variance $\tilde{\nu}_{e,n}^2$.

[7] Introducing such a cost is to match the extensive margin of labor supply, i.e. the employment rate, over the life cycle, as in Karabarbounis (2016).

[8] In the benchmark model, we assume a relatively simple information structure. Alternatively, we could assume that households draw an entry signal about their innate entrepreneurial abilities, providing some information on μ. Assuming a richer information structure would not change our main quantitative results.

Using Bayes' rule and the assumptions of normal densities, one can write how the belief evolves as follows:

$$\tilde{\nu}_{e,n}^2 = \begin{cases} \frac{\nu_e^2\sigma_e^2}{n\nu_e^2+\sigma_e^2} & \text{if} \quad o = E \\ \tilde{\nu}_{e,n-1} & \text{otherwise} \end{cases} \tag{2}$$

$$\tilde{\mu}_{e,n} = \begin{cases} \tilde{\nu}_{e,n}^2\left(\frac{\tilde{\mu}_{e,n-1}}{\tilde{\nu}_{e,n-1}^2} + \frac{\epsilon_{e,n}}{\sigma_e^2}\right) & \text{if} \quad o = E \\ \tilde{\mu}_{e,n-1} & \text{otherwise} \end{cases} \tag{3}$$

where n is a sufficient statistic for computing the posterior variance $\tilde{\nu}_{e,n}^2$. Conditioned on other factors, individual households obtain higher precision about innate abilities as n increases. As shown in Equation (2), given n, $\tilde{\nu}_{e,n}^2$ is increasing in both the variance of the innate productivity shocks ν_e^2 and the variance of the transitory i.i.d. shocks σ_e^2. That is, the *absolute sizes* of both innate productivity and transitory shocks jointly determine the precision of the beliefs given n.

Equation (3) further shows that the posterior mean $\tilde{\mu}_{e,n}$ is a weighted average of the prior mean $\tilde{\mu}_{e,n-1}$ and the signal $\epsilon_{e,n}$, each weighted by the prior variance $\tilde{\nu}_{e,n-1}^2$ and the variance of the transitory shocks σ_e^2. As σ_e^2 increases relative to the prior variance, individuals put a lower weight on the most recent productivity realization—the signal $\epsilon_{e,n}$. That is, the *relative size* of transitory shocks to innate productivity determines the learning speed.[9]

2.4 Income Processes

Wage Income Process Individuals of age j receive a wage income $y_{w,j}$ that is additive in the general equilibrium efficiency wage ω, an exogenous age-dependent component θ_j, a permanent productivity χ_w, and a persistent idiosyncratic wage income shock $\epsilon_{w,j}$:

$$\log y_{w,j} = \log\omega + \log\theta_j + \log\chi_w + \log\epsilon_{w,j}.$$

We assume the ability as workers (both χ_w and ϵ_w) are independent of innate entrepreneurial ability μ.[10]

[9]We could also see that the learning speed is determined by the relative size of the two variances ν_e^2 and σ_e^2 through the formula for $\nu_{e,n}^2$ in Equation (2). With a larger σ_e^2 relative to ν_e^2, the learning speed would be slower since the variance of the posterior belief distribution declines more slowly as n increases.

[10]Alternatively, we can assume a positive correlation between worker and entrepreneurial productivity, reflecting cross-learning between the two occupations or the necessity for entrepreneurs to gain human capital through work experience or education before starting their own firm (Hincapié, 2020; Queiró, 2022). Our primary findings detailed in Sections 4 and 5 remain qualitatively valid.

Entrepreneurial Production and Business Income At the beginning of each period, after observing the signal of innate entrepreneurial productivity ϵ_e, given market prices, entrepreneurs make decisions on how much capital k to rent and how much labor n_b to hire for production. They gain access to a decreasing returns to scale technology $e^{\epsilon_e} f(k, n) = e^{\epsilon_e} (k^\alpha n_b^{1-\alpha})^\eta$, where $\eta < 1$ is the span-of-control parameter. A share η of output goes to factor inputs. Out of this, a fraction of α is going to capital and $1 - \alpha$ is going to labor.

Normalizing the entrepreneurial output price to be one, business income is calculated as revenue net labor and capital rental costs:

$$\begin{aligned} \pi_e(a, \epsilon_e) = & \max_{k, n_b} \{e^{\epsilon_e} f(k, n_b) - \omega n_b - (r + \delta)k - \kappa_o\} \\ s.t. \quad & 0 \leq k \leq \lambda a, \;\; n_b \geq 0, \end{aligned} \tag{4}$$

where r is the deposit rate, δ is the depreciation rate of capital, and κ_o is the fixed operational cost. Note that since the choices for both labor and capital inputs are made after the realization of productivity shocks ϵ_e, business income in terms of profits would always be non-negative if κ_o was not introduced. Thus, we introduce κ_o to generate business loss, as observed in the data. To allow for the impact of borrowing constraints on decisions to become an entrepreneur, we assume that entrepreneurs' capital rental k is limited by a multiple of the collateral, i.e., $k \leq \lambda a$, as in Moll (2014).

Albeit simple, our information structure assumption generates income patterns consistent with recent empirical evidence on the risky nature of entrepreneurial activities.[11]

2.5 Asset Market and Borrowing Constraints

Households have access to competitive financial intermediaries, who receive deposits from both workers and entrepreneurs and rent out capital to entrepreneurs. We focus on within-period borrowing, or capital rental for production purposes. We do not allow borrowing for inter-temporal consumption smoothing, which translates into non-negative financial wealth (i.e., $a \geq 0$). The zero-profit condition of the intermediaries implies a capital rental rate of $r + \delta$ where r is the deposit rate and δ is the depreciation rate of capital.

[11]Using tax return data, DeBacker, Panousi, and Ramnath (2023) show that variations in entrepreneurs' business incomes are mainly due to transitory shocks. In our framework, business incomes are partly random due to the i.i.d. shock ε_n, but also persistent since high-ability types are more likely to receive good signals and stay.

2.6 Corporate Sector

As in Quadrini (2000) and Cagetti and De Nardi (2006), we model a second sector of production populated by a large number of homogeneous firms that we refer to as the non-entrepreneurial corporate sector. Corporate sector firms are not managed by households and operate a constant returns to scale production technology $A_C F(K_C, N_C) = A_C K_C^{\xi} N_C^{1-\xi}$, where A_C is the time-invariant corporate productivity, which can be normalized to one. The terms K_C, N_C are corporate capital and labor demand, respectively. Outputs produced by corporate and entrepreneurial sectors are perfect substitutes. The capital depreciation rate is the same in both sectors. The problem of the corporate sector is thus given by

$$\pi_C = \max_{K_C, N_C \geq 0} \{A_C F_C(K_C, N_C) - \omega N_C - (r + \delta) K_C\}. \tag{5}$$

Factor prices r and ω are thus equated to the marginal productivities as $r = A_c F_K - \delta$ and $\omega = A_c F_N$.

2.7 Government

The government in our model (meant to stand in for all levels—federal, state, and local—in the real world) consumes resources, collects tax revenues, and operates a social security system. The government finances an exogenously given expenditure G with consumption and personal income taxes. Consumption taxes are proportional at rate τ_c. Personal income tax schedule $T^o(\cdot), o \in \{W, E\}$ is a function of pre-government income, which equals the sum of wage income(or business income) and asset income. That is,

$$y^o(a, \epsilon_w, \epsilon_e) = \begin{cases} \omega \chi_w \theta \epsilon_w (1 - l) + ra & \text{for } o = W \\ \pi_e(a, \epsilon_e) + ra & \text{for } o = E \end{cases} \tag{6}$$

The government also operates a balanced pay-as-you-go social security system. Individuals receive social security benefits z that are independent of their contributions and are financed by social security tax τ_{ss}, which is linear in $y^o, o \in \{W, E\}$. The linear tax rate τ_{ss} is exogenously given.

2.8 Recursive Problems

Value of Retirement $(J^V \leq j \leq J)$ Individual households can claim social security as early as age J^V. The value of retirement covers both the voluntary and mandatory retirement ages. Once an individual chooses to retire, that individual cannot return to the labor market in the future. Retired individuals only make the consumption-savings decision and enjoy leisure of unit 1:

$$\begin{aligned} V_j^R(a) = &\max_{a'}\{u(c,1) + \beta[(1-\psi_j)V_{j+1}^R(a') + \psi_j \mathcal{V}(a')]\} \\ s.t. \quad & a' + c(1+\tau_c) = a(1+r) + z \\ & a' \geq \underline{a}. \end{aligned} \tag{P1}$$

Value in Normal Working Age $(0 < j < J^V)$ During normal working ages, households make occupational choice between being a worker or entrepreneur. For $o \in \{W, E\}$,

$$\begin{aligned} V_j^o(\mathbf{x}_j) = &\max_{a',c,l}\{u(c,l;x_e) + \beta[(1-\psi_j)\max_{o'\in\{W,E\}}\{\mathbb{E}V_{j+1}^W(\mathbf{x}_{j+1}), \mathbb{E}V_{j+1}^E(\mathbf{x}_{j+1})\} + \psi_j\mathcal{V}(a')]\} \\ s.t. \quad & a' + c(1+\tau_c) = a(1+r) + (1-\tau_{ss})y_j^o(a,\epsilon_w,\epsilon_e) - T^o(y_j^o + ra) - \kappa_e \mathbb{1}_{\{o\neq E \;\&\; o'=E\}} \\ & \tilde{\mu}_e', \tilde{\nu}_e' = \begin{cases} \Pi(\tilde{\mu}_e', \tilde{\nu}_e' | \tilde{\mu}_e, \tilde{\nu}_e, \epsilon_e) & \text{for } o = E \\ \tilde{\mu}_e, \tilde{\nu}_e & \text{otherwise} \end{cases} \\ & a' \geq \underline{a}, \end{aligned} \tag{P2}$$

where κ_e is the fixed entry cost, and $\mathbb{1}_{\{o\neq E \;\&\; o'=E\}}$ is the indicator function denoting that only households who switch occupations from non-entrepreneurs to entrepreneurs pay such a cost.

Entrepreneurs can exit their businesses either endogenously or exogenously. Endogenous exit occurs when an individual household chooses to switch occupations from entrepreneur to non-entrepreneur. We also incorporate an exogenous separation shock, δ_e, which exclusively applies to incumbent entrepreneurs. If $o = E$ in the current period, the household faces a probability of δ_e to exit entrepreneurship and become a worker in the next period. If the entrepreneur survives this separation shock (with probability $(1-\delta_e)$), she chooses between becoming a worker and continuing being an entrepreneur as described in Problem (P2). The introduction of δ_e serves quantitative purposes, which will be discussed in Section 3. Full details of the recursive problems can be found in Problems (PA1) and (PA2) in Appendix B.1 of the Online Appendix.

Problem (P2) illustrates that in a given period, a household with occupation o makes decisions on assets and occupation for the next period based on idiosyncratic states. If the individual is an entrepreneur in the current period, she uses signal ϵ_e to update her beliefs about her innate entrepreneurial productivity. Workers do not receive entrepreneurial productivity signals and maintain the same beliefs as at the end of the previous period.

Value of Non-retirement in Voluntary Retirement Age $(J^V \leq j < J^R)$ Starting from J^V, individuals can opt for retirement and permanently exit the labor force. Households form expectations by comparing the value of retirement from Problem (P1) and the value of continuing to work. The only distinction between Problem (P2)—the problem during normal working ages—and this stage is the additional option to retire. The recursive problem is formulated as follows.

$$\begin{aligned} V_j^o(\mathbf{x}_j) = \max_{a',c,l}\{u(c,l;x_e) \\ +\beta[(1-\psi_j)\max_{o'\in\{W,E,R\}}\{\mathbb{E}V_{j+1}^W(\mathbf{x}_{j+1}), \mathbb{E}V_{j+1}^E(\mathbf{x}_{j+1}), V_{j+1}^R(a')\} + \psi_j\mathcal{V}(a')]\} \end{aligned} \tag{P3}$$

subject to the same constraints in problem (P2).

2.9 Stationary Competitive Equilibrium

An individual with age j is indexed by states $\mathbf{x}_j = (x_e,\ \chi_w,\ a_j,\ \epsilon_{w,j},\ \tilde{\mu}_{e,j},\ \tilde{\nu}_{e,j},\ \epsilon_{e,j})$. Given a tax structure $\{\tau_c, T^\omega(\cdot), T^b(\cdot), \tau_{ss}\}$ and initial distributions of workers and entrepreneurs over individual states $\{\Gamma_0^W(\mathbf{x}_0), \Gamma_0^E(\mathbf{x}_0)\}$, a **stationary recursive competitive equilibrium** comprises

- prices $\{w, r\}$ and social security benefits z
- policy and value functions for workers and entrepreneurs
- capital and labor demand of the corporate sector
- distribution of households over idiosyncratic states for both workers and entrepreneurs

such that

1. Given prices, the tax structure, and social security benefits, the policy functions solve individual households' problems (P1), (P2), and (P3);
2. The factors demand of the corporate sector solve Equation (5);
3. Capital market, labor market, and social security system are cleared;
4. The government budget is balanced;
5. The distribution of households is stationary.

The equilibrium concept is standard and fully detailed in the Online Appendix.

3 Mapping the Model to the Data

In this section, we outline the process of mapping the model to the data and parameterizing the model in a stationary equilibrium. We employ the simulated method of moments to estimate the model, matching it to data from the US economy in the mid-1990s, in accordance with the availability of several data sources utilized in this paper. The two primary data sources are: (i) the Panel Study of Income Dynamics (PSID) and (ii) the Panel Study of Entrepreneurial Dynamics (PSED).

3.1 Data Sources

PSID Following Heathcote, Perri, and Violante (2010), we focus on the Survey Research Center sample (SRC) and choose a sample of household heads from 1970 to 1997 (corresponding to true years 1969-1996) that includes information on gender, income, age, wealth, self-employment status, and whether the head of a household owns a business.[12] The sample comprises household heads from ages 21 to 75. We include population above the normal retirement age of 65 to take into account the non-trivial fraction of people who are entrepreneurs at older ages. We use the PSID sample to obtain three sets of moments: (i) the entry rate, exit rate, and entrepreneurs as a share of households over the life cycle; (ii) age profiles of assets and earnings; and (iii) personal income tax liabilities.

PSED The PSED investigates the new business start-up process based on nationally representative samples of nascent entrepreneurs (NE) who are active in business creation. We focus on PSED Wave I, which began with screening in 1998-2000 with three follow-up interviews. A control group (CG) of individuals not involved in creating new businesses allows for comparison with NE. The dataset provides valuable information on business creation, including business status, capital structure, legal form, expectations, and performance in terms of sales and employment. Additionally, it offers demographic, labor market experience, and personality traits data for both NE and CG participants. We use the subjective belief data from the PSED sample to document empirical facts on entrepreneurial productivity learning, and the personality traits data to discipline non-pecuniary motives of being an entrepreneur.

Definitions of Entrepreneurs Following Quadrini (2000), we define entrepreneurs in the PSID as *self-employed household heads who are business owners*. For the PSED, we

[12] The entry and exit rates of entrepreneurs at an annual frequency are only available between 1970 and 1997.

limit the observations of *survey-defined NEs* to those who *actually operate and produce*, constituting our final sample of entrepreneurs.[13] Regarding legal forms, the majority of entrepreneurs in both the PSID and PSED are *pass-through business owners*, subject to the personal income tax schedule.[14] Hence, we abstract our analysis from incorporated business owners and only focus on taxation at the entity level.[15]

3.2 Empirical Evidence on Entrepreneurial Learning

Survey Questions on NEs' Expectations We use the PSED Wave I, spanning from 1998 to 2004, which surveys a sample of NEs across four waves. To construct variables on learning, we rely on survey questions regarding NEs' expectations for their businesses' future performance. Following Altig, Barrero, Bloom, Davis, Meyer, and Parker (2022), we use sales information to measure a business's performance.[16]

In Wave 1 (Year 0), the survey asks respondents about their expected sales for both the first and fifth full years of operation. From Wave 2 to 4 (Years 1-3), respondents report their current year's sales and provide predictions for sales in the fifth full year of operation.We showcase the summary statistics of expected and realized sales across different waves in Table A1 of the Online Appendix.

Measuring Learning Based on the data we have, we denote forecasts on sales made in period 0 for periods 1-5 by $\texttt{ESale}_0^q$ for $q = 1, ..., 5$. We denote realized sales in period 1, 2, 3 by $\texttt{RSale}_s$ for $s = 1, 2, 3$. We denote forecasts on sales made in period 1, 2, 3 for period 5 by $\texttt{ESale}_s^5$ for $s = 1, 2, 3$.

Our main measure of forecast errors is the deviation of the realized sales in period s from an entrepreneur's period 0 forecast on its sales for period s scaled by the sum of these two variables.[17] More specifically, forecast errors $\texttt{FError}_0^s$ in period $s = 1, 2, 3$ are constructed as $\texttt{FError}_0^s = \frac{\texttt{RSale}_s - \texttt{ESale}_0^s}{\texttt{RSale}_s + \texttt{ESale}_0^s}$.

[13] See Appendix A.3 of the Online Appendix for more details on the comparison between the PSID, the PSED, and other important data such as the Survey of Consumer Finances (SCF), which oversamples the rich, and the Internal Revenue Service (IRS).

[14] In the PSED, more than 84% of nascent entrepreneurs are pass-throughs, and in the PSID, around 67% of entrepreneurs are unincorporated, which means the share of pass-throughs should be higher than this number.

[15] See Dyrda and Pugsley (2020) for both data and theory where business owners choose between being pass-throughs and incorporations.

[16] The most ideal measure is revenue, which is not available in the data. Alternatively, we can use employment as a measure. However, approximately half of the NEs opt for "mere" self-employment in the PSED, meaning they have no employees and no intention to expand their employment size.

[17] We do normalization this way to ensure the forecast errors fall within a certain bounded interval, in our case, $[-1, 1]$. Alternatively, we can define forecast error as $log(\texttt{RSale}_s) - log(\texttt{ESale}_0^s)$, which does not change our main empirical findings.

Our main measure of forecast revision is the deviation of the prediction on sales for period 5 made in period s from the prediction on sales for period 5 made in period 0 scaled by the sum of the two variables. More specifically, forecast revision in period s on period-5 performance $\texttt{FRev}_s^5$ for $s = 1, 2, 3$ is constructed as $\texttt{FRev}_s^5 = \frac{\texttt{ESale}_s^5 - \texttt{ESale}_0^5}{\texttt{ESale}_s^5 + \texttt{ESale}_0^5}$.

(a) Distribution of Forecast Errors (b) Forecast Revision Updating

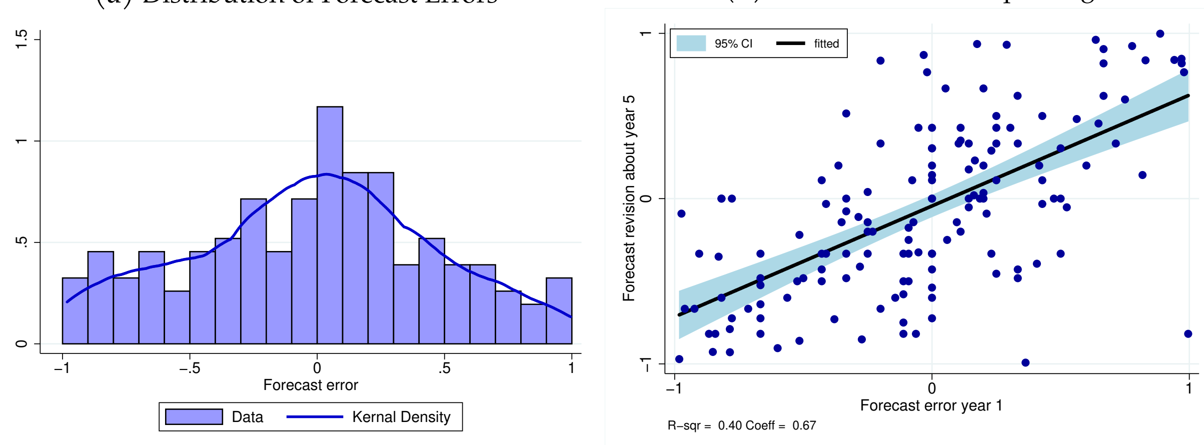

Figure 3: Size and Evolution of Entrepreneurial Uncertainty

Distribution of Forecast Errors Figure 3(a) displays the distribution of forecast errors and their kernel density for Year 1, illustrating the initial uncertainty faced by nascent entrepreneurs. The distribution of `FError` is nearly symmetric in Year 1, with a standard deviation of approximately 0.48. Although some entrepreneurs make fairly accurate forecasts, with `FError` close to 0, over 30% of them predict sales that deviate by more than 50% from actual sales (absolute value of `FError` ≥ 0.5), indicating the substantial uncertainty faced by new entrepreneurs.[18]

Forecast Errors Predict Future Forecast Revisions Figure 3(b) displays a scatter plot of the joint distribution of forecasts on Year-5 sales and realized sales in Year 1, which provides insights into how long-term forecasts are revised after entrepreneurs observe their actual sales. Specifically, we regress the forecast revision for Year-5 sales made in Year 1 on the forecast errors in Year 1. We obtain an R-square value of 0.40, indicating a

[18]We also find no significant difference in initial distributions of forecast errors based on gender, age, education levels, industry, or previous work experience. See Table A2 in Appendix A of the Online Appendix for more details. In particular, that the initial distribution of forecast errors changes little with respect to entrepreneurs' age indicates that these errors likely pertain to uncertainty about entrepreneurial ability rather than other factors such as demand. This evidence further supports our "no cross-learning between occupations" assumption discussed in Section 2.

strong linear relationship between forecast revisions and forecast errors. The coefficient of 0.67 suggests that for every dollar of realized sales above their previous period's forecast, entrepreneurs increase their future sales forecasts by 67 cents. In addition, through the vertical axis of Figure 3(b), the distribution of forecast revisions is very dispersed, suggesting information uncertainty is still large even after learning.[19]

We acknowledge that entrepreneurs learning from forecast errors to revise their forecasts for future sales does not necessarily mean that they make more precise forecasts as the entrepreneurial spell lengthens. Due to the short panel limitations of the PSED, we could not observe whether the accuracy of forecasts improves or not over the duration of an entrepreneurial spell. We supplement our finding in Figure 3(b) with data from the PSID, which shows that the exit rate declines as the duration of entrepreneurial spells increases, as demonstrated in Section 3.4. The fact that the exit rate declines in business duration suggests that entrepreneurs do not merely update their forecasts randomly.

3.3 Functional Specification and Parameterization

We begin with the subset of parameters that can be distinctly identified outside our model including those on demographics, wage income process, production technology, and government policies, then consider those estimated within the model including those on preferences, uncertainty, and entrepreneurial ability learning. We summarize the estimation results in Table 1.

3.3.1 Demographics, Preferences, and Bequests

A model period is equivalent to one year. Individuals are born at age 21 (model period 1) and can voluntarily retire at age 62 (model period $J^V = 42$), retire compulsorily at age 80 (model period $J^R = 60$), and die with certainty at model age 101 (model period $J = 81$). The population growth rate g_n is 0.011 at an annual rate, and the mortality probability is taken from Bell and Miller (2005).

We set $\zeta = 4$, standard in the macro labor literature, and choose β and γ so that the stationary equilibrium of the economy with the benchmark tax system features a capital-output ratio of 2.7 and an average share of time worked of one-third of the time endowment.

[19]We also find that the slope of forecast revision for Year-5 sales made in Year 1 on Year-1 forecast errors is quite robust, as the coefficient changes little after controlling for factors such as gender, age, education, industry, previous work experience, and first-time entrepreneurs. See Table A3 in Appendix A of the Online Appendix for more details.

A unit of time in the model corresponds to 120 total hours per week, which can be allocated between work and leisure. Workers choose weekly working hours from a set of four discrete choices $\{0, 20, 40, 50\}$. Leisure l for workers who work zero hours is equal to one. Workers who work positive hours derive disutility from working hours h and a fixed utility cost of working $\phi_{w,0}$. We choose $\phi_{w,0}$ to match a 70% employment rate in the US for the entire population between ages 21 and 65.

Non-pecuniary Motives We use survey questions from the PSED to demonstrate how personality traits affect the entry into entrepreneurship, which in turn helps discipline the non-pecuniary motives for being an entrepreneur. Following Lise and Postel-Vinay (2020), we use the method of the principal component analysis (PCA) to summarize the original 25 questions on personality traits into six key traits—that is, the Big 5 plus *love of business* (LoB).[20] Among the six personality traits, only LoB is found to be significantly different between the sample of NEs (nascent entrepreneurs) and the CG (control group).[21] The constructed LoB scores are normalized to lie within the $[0, 1]$ range. We present the distribution of the LoB trait scores in Figure 4(a).

The non-pecuniary motive to be an entrepreneur in our model is manifested as a fixed utility cost of being an entrepreneur, $g(x_e)$ in Equation (1), which is a linear function of LoB state x_e. To discipline the distribution of LoB states x_e, we approximate the distribution of the LoB trait scores to a beta distribution with two shape parameters equal to 3.2 and 2.8. Finally, we discretize the beta distribution with seven states for simulation, as shown in Figure 4(a). We specify $g(x_e) = \phi_{e,0} + \phi_{e,1}x_e$. The slope parameter $\phi_{e,1}$ captures the differences in utility costs faced by agents with different LoB states and is set to match the difference in mean LoB scores between entrepreneurs and workers in our PSED sample. The idea is that a higher value of $\phi_{e,1}$ implies a larger variation in utility costs of being an entrepreneur, resulting in a greater difference in mean LoB scores between workers and entrepreneurs. The intercept parameter $\phi_{e,0}$ is set to match the share of entrepreneurs in PSID, analogous to disciplining the fixed utility cost of workers ϕ_w by the employment rate.

Bequests Following De Nardi (2004) and Lockwood (2018), we specify the bequest utility function as $\mathcal{V}(b) = (\frac{\phi_b}{1-\phi_b})^{\tilde{\zeta}} \frac{(\frac{\phi_b}{1-\phi_b}c_b+b)^{1-\tilde{\zeta}}}{1-\zeta}$, where $\tilde{\zeta} = 1-\gamma(\zeta-1)$ captures the weight on consumption that is consistent with flow utility $u(.)$, $c_b > 0$ is the threshold consumption

[20] As emphasized in Hamilton, Papageorge, and Pande (2019), a large literature in psychology uses five traits (the Big 5) to comprehensively describe an individual's personality.

[21] We demonstrate this result in Figure A3 and Table A7 in Appendix A of the Online Appendix.

Table 1: Model Parameterization

Parameter	Description	Value	Source/Target
Demographics			
J^V	Youngest age to claim retirement	42	Age 62
J^R	Age of mandatory retirement	60	Age 80
J	Age of death	81	Age 101
g_n	Population growth rate	0.011	Conesa, Kitao, and Krueger (2009)
$\{\psi_j\}_{j=1,\ldots,81}$	Mortality probability		Bell and Miller (2005)
Preferences			
ζ	Risk aversion	4	IES $= 0.5$
γ	Intensity of consumption	0.40	2,000 annual hours for workers
β	Discount factor	1.00	K/Y $= 2.7$
ϕ_ω	Fixed cost of working	0.25	Employment rate $= 0.70$
c_b	Threshold consumption level	0.30	\$17000
ϕ_b	Marginal propensity to bequeath	0.95	Bequest as a share of wealth $= 0.6$
$(\beta_{e,1}, \beta_{e,2})$	Beta distribution: LoB score	(3.2, 2.8)	PSED LoB distribution
$\phi_{e,0}$	Fixed cost of entrep.: intercept	0.60	Share of entrep. in population $= 9.7\%$
$\phi_{e,1}$	Fixed cost of entrep.: slope	–0.09	Diff. in mean LoB: entrep. to worker $= 0.10$
Entrepreneurial productivity learning process & production			
μ_e	Mean: innate entrep. prod.	1.31	Mean business to wage income $= 2.5$
ν_e	Std. dev.: innate entrep. prod.	0.17	Std. dev. of forecasting error $= 0.40$
σ_e	Std. dev.: i.i.d. shocks	0.29	Slope of forecast revision $= 0.67$
κ_o	Per period operational cost	0.02	Frac. of entrep. with negative income $= 0.05$
κ_e	One-time entry cost	0.05	Annual exit rate $= 0.20$
δ_e	Exogenous separation rate	0.02	Exit rate for duration of 10+ years $= 0.08$
λ	Collateral parameter	1.50	Median wealth entrep. to worker $= 6.0$
Wage income process			
$\{\theta_j\}_{j=1,\ldots,60}$	Age-dependent wage prod.		Hansen (1993)
ρ_w	Wage income shock: persistence	0.98	Conesa, Kitao, and Krueger (2009)
σ_w	Wage income shock: std. dev.	0.17	Conesa, Kitao, and Krueger (2009)
σ_χ	Permanent types dist.: std. dev.	0.37	Conesa, Kitao, and Krueger (2009)
Production technology			
ξ	Capital share: corporate	0.36	Corporate labor share from NIPA
η	Span of control: entrepreneurs	0.79	Buera, Kaboski, and Shin (2011)
δ	Capital depreciation rate	0.06	BEA fixed asset tables
α	Capital share: entrepreneurs	0.36	Same as ξ
Government policies			
τ_c	Consumption tax rate	0.065	Bhandari and McGrattan (2021)
τ_{ss}	Payroll tax rate	0.124	Conesa, Kitao, and Krueger (2009)
κ_0	Personal income tax: level shifter	0.091	PSID estimation
κ_1	Personal income tax: progressivity	0.142	PSID estimation

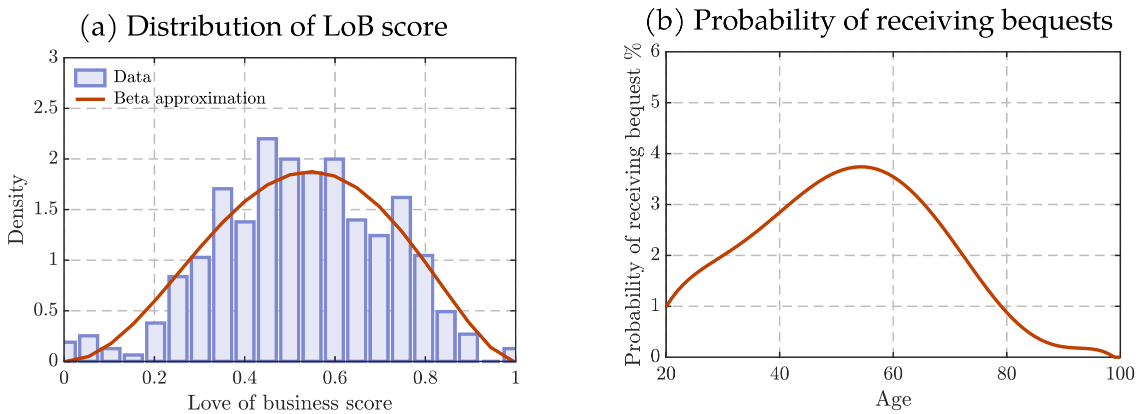

Figure 4: Model Inputs of LoB Score Distribution and Bequests

level to leave a bequest, and $\phi_b \in [0, 1)$ is the marginal propensity to bequeath. We set the c_b to target a threshold level of $\$17000$, as estimated by Lockwood (2018), and calibrate ϕ_b to match the ratio of mean bequests to mean wealth level 60%.

As documented by Cagetti (2003), the probability of receiving bequests is hump-shaped in age, and the ratio of average bequests conditional on receiving to average disposable incomes is roughly constant by age. We directly take the bequest probability from Cagetti (2003), as shown in Figure 4(b), and set the bequest to income ratio to 1.67, the value that balances bequests left and bequests received by agents in equilibrium.

3.3.2 Wages, Production, and Entrepreneurial Productivity Learning

Wage Income Process We take age-productivity profile $\{\theta_j\}_{j=1}^{JR-1}$ from Hansen (1993). Following Conesa, Kitao, and Krueger (2009), we discretize permanent worker productivity into two types with equal probability. We consider $\chi_w \in \{\chi_1, \chi_2\}$, with $\chi_1 = e^{-\sigma_\chi}, \chi_2 = e^{\sigma_\chi}$ such that $E(\log(\chi_w)) = 0, \text{var}(\log(\chi_w)) = \sigma_\chi^2$. Idiosyncratic shocks of wage income follow a simple AR(1) process with persistence parameter ρ_w and unconditional variance σ_w^2:

$$\log \epsilon_{w,j} = \rho_\omega \log \epsilon_{w,j-1} + \varepsilon_{w,j}, \quad \varepsilon_w \sim i.i.d.\mathcal{N}(0, \sigma_w^2)$$

We take the parameters from Conesa, Kitao, and Krueger (2009) and approximate the stochastic process with five discrete states.

Production Technology The capital share of corporate firms' production function ξ is set to be 0.36 to match the labor income share of the corporate sector from the BEA-NIPA. For simplicity, we make the value of the capital share of the entrepreneurial sector equal to that of the corporate sector. The span of control parameter η is set to be 0.79 following Buera, Kaboski, and Shin (2011). Taking the scale of production η into consideration leads to a capital share $\alpha\gamma = 0.28$ for the entrepreneurial sector, which is close to the value used in the macro literature on entrepreneurs (e.g., Buera, Kaboski, and Shin (2011), Cagetti and De Nardi (2006)). The capital depreciation rate is set to be 0.06 based on the estimates using the BEA fixed asset tables. Since individuals' wealth accumulation significantly affects whether they are financially constrained or not once they become an entrepreneur, we discipline the collateral parameter λ to target the PSID moment that the ratio of the median wealth held by entrepreneurs to that held by workers is around six.

Learning Process on Innate Entrepreneurial Ability The two key parameters governing the entrepreneurial learning process are the variance of the distribution of transitory shocks σ_e^2 and the variance of the distribution of innate entrepreneurial ability types ν_e^2. Note that ν_e^2 captures the initial uncertainty size faced by entrant entrepreneurs, which has a direct empirical counterpart—the variance of forecast errors in the initial year, as documented in Figure 3(a). The relative size of σ_e^2 and ν_e^2 determines the learning speed in the model, which can be disciplined by the magnitude of forecast revision updating after observing forecast errors, as documented in Figure 3(b). The mean of the distribution of innate ability types μ_e determines the level of profits earned by entrepreneurs and is chosen to match the ratio of the mean income of entrepreneurs to that of workers. Based on our calibration, the size of the ex post risk σ_e^2, 0.29, is greater than the size of the ex ante risk ν_e^2, 0.17.

It is worth noting that models in the existing literature on learning and firm dynamics, such as Jovanovic (1982), Arkolakis, Papageorgiou, and Timoshenko (2018), and Chen et al. (2023), feature very fast learning dynamics.[22] Our model differs from those firm dynamics models with learning in the feature of occupation choice between workers and entrepreneurs. While learning can be fast for entrepreneurs, it is slower on average since agents aren't always entrepreneurs. As shown in Figure 5, the standard deviation of belief about innate entrepreneurial ability decreases by around 70% within ten years for all-time entrepreneurs, but only by around 30% after 50 years for the average individual.

We discipline the remaining three parameters on entry and exit, namely, (i) the exogenous separation rate δ_e, (ii) the per-period operational cost κ_e, and (iii) the one-time entry cost κ_o as follows. In the data, the exit rate declines during the first 10 years of entrepreneurial operation and plateaus thereafter. In contrast, the model's exit rate would continue to decline (although at a decreasing rate) as the entrepreneurial spell

[22]One exception is Kochen (2022), which introduces the age-specific variances of transitory shocks.

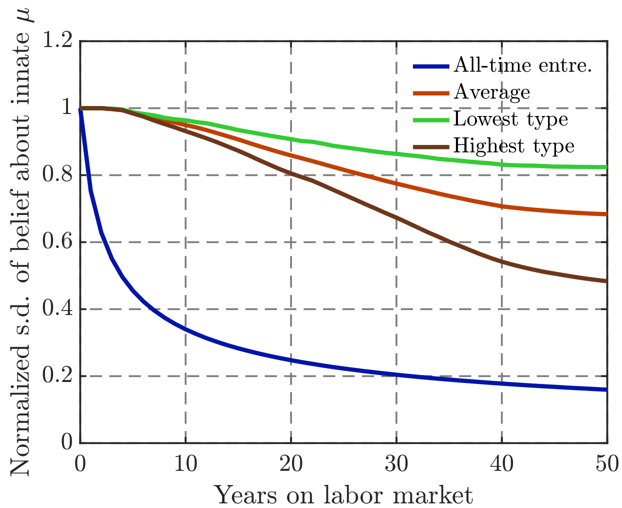

Figure 5: Bayesian Learning Speed Implied by the Benchmark Model

lengthens, if no exogenous separation shocks were introduced. Thus, we choose the exogenous exit rate δ_e to target the exit rate for entrepreneurial spells exceeding 10 years.

We introduce κ_e and κ_o to capture the business losses observed in the data. A higher entry cost κ_e creates a greater entry barrier and a larger effective cost in the first period of business operation, leading to lower exit rates in subsequent periods and a steeper slope of the fraction of entrepreneurs with negative incomes by age. We pick κ_e and κ_o to match the cross-sectional annual exit rate and the fraction of entrepreneurs making negative incomes.

3.3.3 Government Policies

Following Bhandari and McGrattan (2021), we set the consumption tax rate τ_c to 0.065. We set the payroll tax rate τ_{ss} to 0.124, following Conesa, Kitao, and Krueger (2009). The benefit z is determined by the balanced government budget in equilibrium.

Given that logged after-tax income and logged pre-tax income exhibit approximately a linear relationship in US data, Benabou (2002) and Heathcote, Storesletten, and Violante (2017) approximate the progressive tax system with the following non-linear function:

$$T(y) = y - (1 - \kappa_0)(y)^{(1-\kappa_1)} \tag{7}$$

Alternatively,

$$\ln(y - T(y)) = \ln(1 - \kappa_0) + (1 - \kappa_1) \ln y \tag{8}$$

where y represents the pre-government income, as defined in Equation (6). The expression $T(y)$ denotes the associated tax liabilities, and $y - T(y)$ is the post-government income. Equation (7) characterizes the tax function with a level parameter κ_0 and a progressivity parameter κ_1. A tax schedule with $\kappa_1 = 0$ corresponds to a proportional income tax system. As κ_1 increases, the tax system becomes more progressive.

Following Heathcote, Storesletten, and Violante (2017), we recover parameters κ_0 and κ_1 from our PSID sample using an ordinary least squares (OLS) regression based on Equation (8). We obtain tax liabilities by submitting income and demographic variables from the PSID to the NBER TAXSIM program to calculate federal and state-level income taxes, as well as deductions.[23]. In our benchmark estimation, we pool entrepreneurs and workers together and measure pre-tax incomes y as the sum of labor income, self-employment income, and asset income. We obtain $\kappa_0 = 0.0912$ and $\kappa_1 = 0.1416$, and we plot the tax schedule in Figure 10, which is represented by the blue solid curve.[24]

3.4 Model Performance

In the previous section, we discipline the entrepreneurial learning process using the PSED data on entrepreneurial forecasts updating. However, due to the limitation of PSED being a short panel where the evolution of forecast precision cannot be observed, we supplement it with the PSID, which shows a decline in exit rates with respect to entrepreneurial spell duration. In this section, we demonstrate that our model well replicates the life cycle outcomes of entrepreneurs, especially the entry and exit patterns, which serves as indirect evidence on learning and experimentation in entrepreneurship.

Entrepreneurship over the Life Cycle Using the PSID, we observe strong life cycle patterns of business entry and exit. Even though none of the moments are directly targeted, our model successfully replicates the entry, exit, and the overall share of entrepreneurs over the life cycle, as shown in Figure 6. The entry rate peaks at around age 45-50 at 3.0%, and declines thereafter. The exit rate decreases during working age and increases after age 60 as a result of retirement.

Since younger individuals are more likely to be first-time entrepreneurs and have a shorter entrepreneurial spell, the higher exit probability of young entrepreneurs shown in Figure 6 aligns with the declining exit rate by entrepreneurial spell documented in Figure 7(a).

This outcome stems from the interplay between two key elements of the model: asset

[23] More details about TAXSIM can be found in Feenberg and Coutts (1993).

[24] Our benchmark tax schedule should thus be interpreted as statutory. Our main results established in Section 5 will not be affected if we use the effective rates estimated in Bhandari and McGrattan (2021) instead.

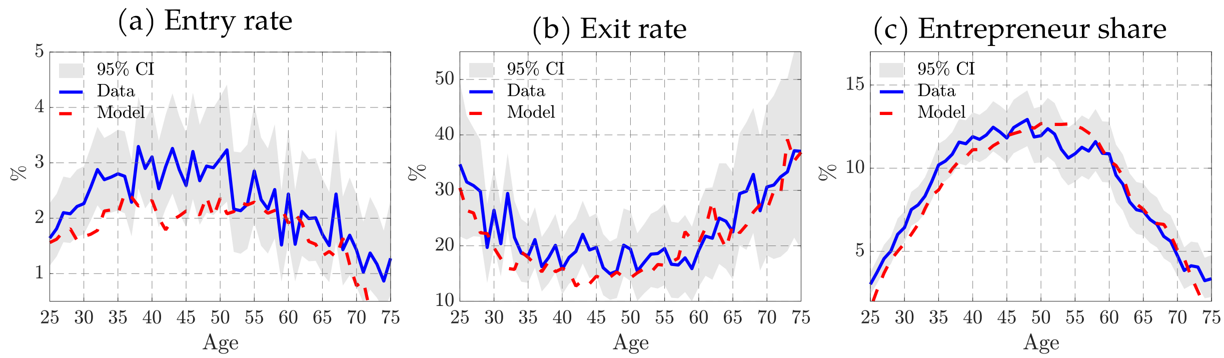

Figure 6: Entry and Exit Rates over the Life Cycle: Model v.s. Data

Note:(i) the entry rate is calculated as the number of entrants as a fraction of non-entrepreneurs in the previous period; and (ii) the exit rate is measured as the number of exit entrepreneurs as a fraction of entrepreneurs in the previous period.

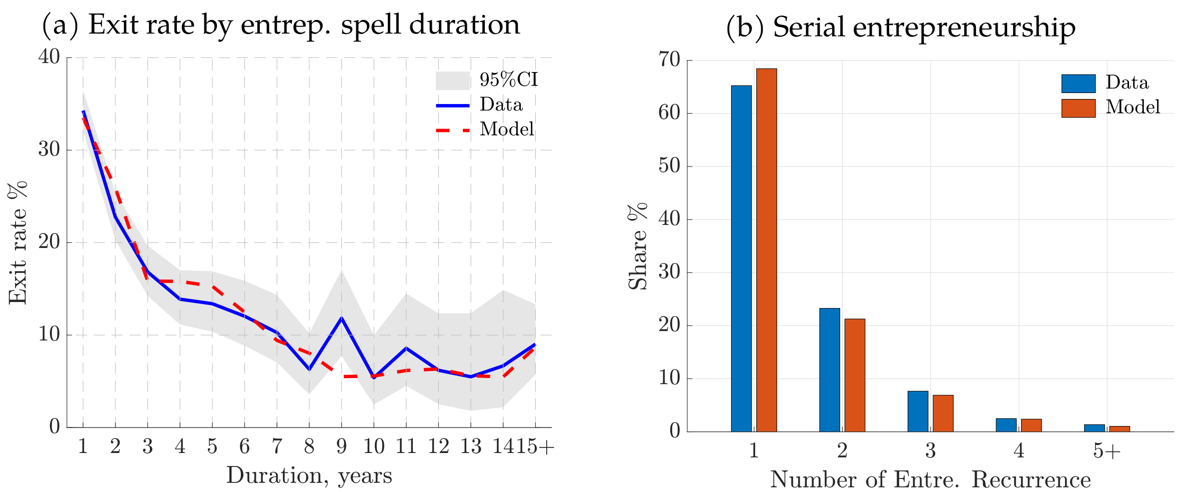

Figure 7: Model Fit: Exit by Duration and Serial Entrepreneurship

Note: Panel (a): the horizontal axis represents the number of years a household head has been an entrepreneur, while the vertical axis indicates the exit rate. Panel (b): the horizontal axis represents the number of entrepreneurial recurrences; for example, one indicates that an individual has been an entrepreneur once during their lifetime in our PSID sample. The vertical axis represents the share of a specific group of individuals on the horizontal axis among all individuals who have been entrepreneurs at least once.

accumulation subject to financial frictions and reduced uncertainty through learning. The earlier the uncertainty is resolved, the longer an entrepreneur can expect to operate, thereby reaping greater earnings if the entrepreneur is innately productive. Consequently, agents are inclined to experiment with entrepreneurship to learn about their innate abilities as early as possible. However, the asset accumulation channel delays young agents' entry decisions. On the one hand, young agents with low asset levels cannot insure themselves against a low productivity shock. Even with a high shock realization, they are unable to scale up production to increase earnings as a result of collateral constraints. Therefore, the asset and information channels jointly determine that the entry age of entrepreneurs peaks in middle age, consistent with the data shown in Figure 6.

The declining exit rate for working-age individuals is primarily driven by reduced uncertainty through learning. Young agents, possessing little information about their innate entrepreneurial productivity, experiment with entrepreneurship and consequently exit with a high probability. Only those who discover themselves innately productive will ultimately stay. This selection through the learning mechanism is also evident when we condition the exit rate by duration, as demonstrated in Figure 7(a). With re-calibration under perfect information case, where the learning channel driving entrepreneurial choice is absent, entrepreneurial exits are primarily explained by the exogenous separation shock δ_e. Consequently, we do not see the pattern of declining exit rates with respect to working age, especially at young ages.[25]

The overall share of entrepreneurs follows a hump-shaped pattern, peaking at middle age, which results from both entry and exit dynamics. This finding is broadly consistent with empirical evidence in Azoulay, Jones, Kim, and Miranda (2020), which shows that successful entrepreneurs are middle-aged using IRS K-1 and Census Bureau business data.

We find these patterns to be robust regardless of whether we use survey data (e.g., PSID, SCF, Current Population Survey (CPS)) or administrative data.[26] For instance, Bhandari et al. (2023) uses US administrative data from the IRS and discovers similar entry and exit patterns, even when applying slightly different definitions of entrepreneurs.

One potential concern is that when an entrepreneur chooses to exit and become a worker, they may forget the learned information about their innate entrepreneurial productivity, whereas our model assumes that agents can retain this information perfectly. We examine the distribution of recurrent entrepreneurial activities in Figure 7(b) and observe that more than 65% of household heads in the data have been entrepreneurs only once during their lifetime. Our model replicates the data well, confirming that our assumption is not extreme.

Other Key Moments Although we use relatively direct evidence on entrepreneurs' expectation formation to discipline the learning process instead of indirect moments

[25] For more details, see Section B.4.1 in Appendix B of the Online Appendix.

[26] See Figure A6 in Appendix A of the Online Appendix for patterns using monthly CPS panel data, which is consistent with empirical findings in Evans and Leighton (1989). Additionally, refer to Figure A8 for the humped-shaped nature of participation into entrepreneurship using the SCF.

such as the age-profile of entrepreneurial income, as in Dillon and Stanton (2018) and Hincapié (2020), our calibrated model can well replicate the mean and standard deviation of entrepreneurial income with respect to the entrepreneurial spell. Moreover, our model also well captures both the income and wealth distributions observed in the PSID and the SCF, as in Cagetti and De Nardi (2006). Notably, it accurately replicates the fraction of entrepreneurs within each wealth quantile of the overall wealth distribution, including the top 1%. The results, along with other model fit moments, are detailed in Appendix B.3 of the Online Appendix.

4 The Value of Learning and the Cost of Uncertainty

4.1 The Value of Learning

We measure the value of learning in terms of three objective moments: (1) aggregate entrepreneur share across ages; (2) discounted lifetime business income; and (3) discounted lifetime total income (i.e. the sum of wage income, business income, and asset income).

The exercise consists of checking the deviation in objective moments from the benchmark economy if we do not allow agents to update their beliefs about innate ability at a specific age while still permitting them to update beliefs at other ages.

As we can see from Figure 8, when agents know that they are not able to learn at a certain age, they are less likely to be an entrepreneur, even at older ages, compared to the benchmark economy. Consequently, the discounted lifetime business income also becomes less for all age groups. So does the discounted lifetime total income due to worsened occupation allocation. This implies that there is always a positive value of learning about innate ability and reducing the uncertainty associated with it as well.

Overall, the value of learning is monotonically decreasing in age except for the very young in terms of the deviation of entrepreneur share. The reason is that young agents, who possess a small amount of assets, do not gain as much from learning since their earnings are constrained by financial frictions even if they find themselves innately productive, and they lack enough assets to smooth consumption with low productivity shock realizations. As they age, the financial constraint is gradually relaxed and the horizon effect dominates so that the value of learning is strictly decreasing in age.

The value of learning, in terms of aggregate entrepreneur share, peaks between ages 30 and 34. This implies that if agents ages 30 to 34 do not learn about their innate productivity, the aggregate entrepreneur share across ages would decrease from the

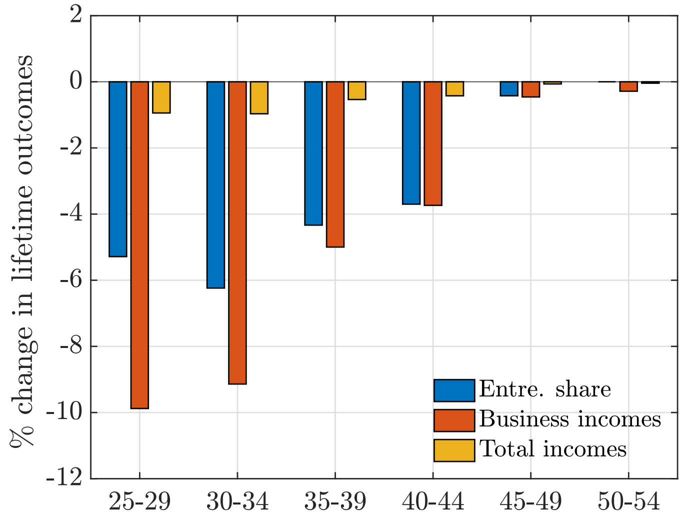

Figure 8: Value of Learning by Age

benchmark 9% to the counterfactual 8.5%. In other words, the timing of becoming an entrepreneur makes a significant difference. The earlier an individual becomes an entrepreneur to learn about their innate type, the sooner the uncertainty can be resolved. Consequently, this leads to better occupational sorting, a higher aggregate entrepreneur share, and increased lifetime income in the cross section.

4.2 The Cost of Uncertainty

To quantify the cost of uncertainty arising from imperfect information in the sense that agents do not know their innate entrepreneurial ability upon entering the labor market, we compare our benchmark economy with the case of perfect information. In the perfect information scenario, people already know their true entrepreneurial ability before entering the labor market. After they decide to be an entrepreneur, there will still be transitory shocks realized to their innate productivity, which are essentially the productivity given which the output is produced. Thus, the case of perfect information can be easily nested by our benchmark model.

We examine the lifetime outcomes by innate entrepreneurial ability types and report the results in Table 2. We consider seven productivity levels, ranging from -3 to +3 standard deviations from the mean. For each type, we analyze three lifetime variables: (1) entrepreneur share; (2) the share of business income in total income; and (3) total income, where the median type's income is normalized to one. We then compare two cases.

Two messages are almost immediate. First, in both cases, the likelihood of en-

Table 2: Lifetime Outcomes by Innate Entrepreneur Ability Types

Innate ability types	-3 sd	-2 sd	-1 sd	0 sd	+1 sd	+2 sd	+3 sd
Panel (A). Benchmark with informational frictions and learning							
Lifetime entrepreneur share	0.01	0.01	0.02	0.04	0.14	0.34	0.39
Lifetime y^b in total y	0.00	0.00	0.01	0.02	0.12	0.40	0.61
Lifetime incomes (normalized)	1.00	1.00	1.00	1.00	1.06	1.35	1.87
Panel (B). Perfect information							
Lifetime entrepreneur share	0.00	0.00	0.00	0.00	0.12	0.71	0.94
Lifetime y^b in total y	0.00	0.00	0.00	0.00	0.09	0.64	0.99
Lifetime incomes(normalized)	1.00	1.00	1.00	1.00	1.04	1.48	2.56

Note: Lifetime incomes of all entrepreneurial ability types are normalized by the average income of the median type (0 sd).

trepreneurship, business income share, and total income increase with entrepreneurial ability. Second, when comparing the cases, under perfect information, only individuals with high innate entrepreneurial productivity become entrepreneurs in equilibrium, while in benchmark scenario where information is unclear, even low-ability types have some chance to become entrepreneurs.

Furthermore, we can gauge the cost of informational frictions by comparing our benchmark case with the perfect information case. Switching to the case without informational frictions significantly improves occupational allocation by increasing the chance for agents with high entrepreneurial productivity (above the mean) to become entrepreneurs during their lifetime, as well as the share of business income in their total incomes. For instance, with perfect information, the likelihood of individuals with the highest type (+3 standard deviations) becoming entrepreneurs during their lifetime rises from 39% to 94%, with 99% of their income derived from business rather than wage income. This result also suggests that the value of learning is higher for agents with higher entrepreneurial productivity.

4.3 Interaction between Informational and Financial Frictions

Both entrepreneurial productivity learning in the presence of informational frictions and asset accumulation in the presence of financial frictions have an impact on entrepreneurial choice. To better understand the role of learning in shaping entrepreneurship over the life cycle, it is crucial to consider the significance of the asset accumulation channel and how it interacts with informational frictions.

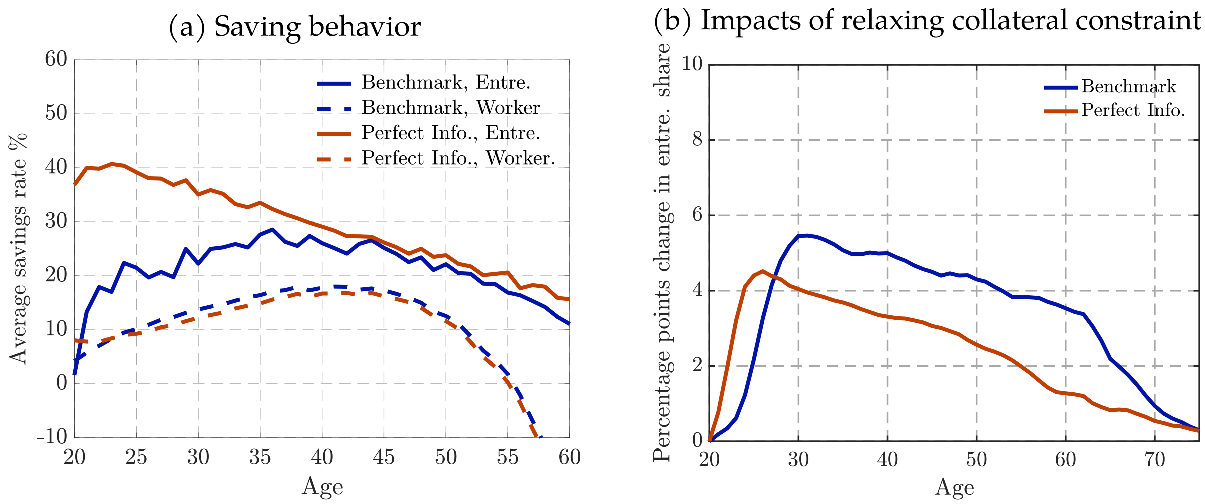

Figure 9: Interaction between Informational Frictions and Financial Frictions.

Note: Panel (a): a plot of average saving rates for entrepreneurs and workers under cases of benchmark and perfect information. The savings rate is defined as the growth in net assets during a period, divided by the total income in that same period. Panel (b): a plot of the percentage point change in entrepreneur share before and after the collateral constraint is relaxed.

To investigate the interaction, we first examine the saving behavior for workers and entrepreneurs in both benchmark and perfect information scenarios. As depicted in Figure 9(a), workers exhibit highly similar hump-shaped saving patterns throughout their life cycle in both cases. However, entrepreneurs demonstrate significant differences between the two scenarios. Under perfect information, high-entrepreneurial-ability agents save considerably from the outset. In contrast, when faced with limited information on their abilities (as in the benchmark scenario), individuals become entrepreneurs to learn more about their productivities. However, uncertain about their future entrepreneurial pursuits, they save substantially less during the early stages of their careers compared to the perfect information case. This finding suggests that in the presence of substantial uncertainty due to imperfect information, the self-financing mechanism, as emphasized by Moll (2014), will be hindered given the persistence of idiosyncratic shocks.[27]

Next, we relax the collateral constraint faced by entrepreneurs (as defined in Problem (4)) by increasing the value of the collateral parameter, λ, from the calibrated 1.5 to 2.0. This means that entrepreneurs are now able to borrow up to 100% of their own assets to finance capital rental, as opposed to the previous 50%. The rationale behind this exercise is that the collateral requirement, λ, is a key parameter that greatly influences the saving motive and the severity of financial frictions.

[27] The saving behavior by entrepreneurial ability types over the life cycle in both benchmark and perfect information cases can be found in Figure A14 in Appendix B of the Online Appendix.

The results, shown in Figure 9(b), reveal that in the benchmark case, relaxing the borrowing constraint does not substantially increase the entrepreneur share among very young people (age 20-25). However, under perfect information, where only the asset accumulation channel is present, the share of young entrepreneurs rises immediately and significantly. This is because, under perfect information, the collateral constraint is only binding for entrepreneurs with high innate ability. Since agents are perfectly aware of their abilities, once the collateral constraint is relaxed, those high-ability individuals will enter immediately. However, when entrepreneurial productivity is unknown, individuals may choose to accumulate additional wealth before entering (also as suggested by Panel (a) they do not save as much as in the perfect information case at young ages), in case of a low realization of productivity shocks. The overall effects of easing collateral constraints are larger in the benchmark scenario with uncertainty about innate ability. This is due to the higher number of marginal entrepreneurs present compared to the perfect information case.[28]

5 Policy Experiments

Our model's key implications from the previous section suggest that policies aimed at reviving entrepreneurship should prioritize targeting young individuals. By enabling earlier entry into entrepreneurship, these policies disproportionately benefit those with high innate entrepreneurial abilities. These policies not only boost entrepreneurship but also improve welfare because more output is produced by highly productive entrepreneurs under such policy reforms. To illustrate this idea, we consider two types of policies that have been widely proposed or even implemented in practice and gauge their impacts on entrepreneurship.

The first policy experiment is to directly subsidize entrepreneurs based on observed characteristics, particularly age, considering either entrant entrepreneurs or all operating entrepreneurs.[29]

Another fiscal policy instrument that is commonly used by the government to influence self-employed business owners is the personal income tax. We focus on the progressivity of the tax system since it directly interacts with the uncertainty and risks of entrepreneurial incomes. A more progressive tax system trades off insurance for low-income realizations with compressed returns for high-income realizations, achieving

[28]See Table A16 in Appendix B of the Online Appendix for more details.

[29]For example, the Singapore government provides a start-up capital grant of 50,000 SG dollars to first-time entrepreneurs with innovative ideas.

effects similar to an age-dependent tax favoring the young, who typically have lower incomes and larger demand for insurance due to uncertainty and limited wealth.[30] Thus, in the second policy experiment, we evaluate the impact of the US progressive personal income tax scheme on entrepreneurship by comparing it to a counterfactual flat tax scheme with zero progressivity that generates the same tax revenue.

Throughout all counterfactual experiments, we focus on steady-state comparisons and apply the policy changes exclusively to the entrepreneurial sector.

5.1 Subsidizing Entrepreneurship

In this section, we explore various types of subsidies for business owners based on their characteristics: all operating business owners, business entrants, and business owners and entrants at different ages. To ensure comparability across experiments, we set the subsidy level for each experiment such that the total amount of subsidy remains the same across all experiments. Following Bhandari and McGrattan (2021), we target the size of the total subsidy to be 4% of business incomes before the reform, a figure motivated by the tax relief provided to pass-through business owners under the 2017 US Tax Cuts and Jobs Act reform.

Table 3 presents the effects of subsidy policy experiments on the size and quality of entrepreneurship. Panel (A) displays the impacts of experiments targeting business entrants only, while Panel (B) reveals the impacts of experiments targeting all operating business owners.

Column (1a) demonstrates that an 18% subsidy of the median income for entrants ages 25-34 leads to a 4.19 percentage point increase in the entrepreneur population share compared to the benchmark model. Among all innate productivity types, the highest type experiences a disproportionately larger increase.

Columns (2a) to (3a) display results for targeting entrants ages 35-44 and 45-54, respectively. As observed in Section 4, the value of resolving uncertainty through learning decreases with age. Consequently, the increase in the entrepreneur population share and output both diminish as the targeted age group rises. Furthermore, the older the age group, the smaller the differences in the entrepreneur share increase between low-ability and high-ability groups.

Finally, in column (4a), we remove the age restriction by assuming that the government subsidizes entrants at all life stages (i.e., ages 20-80). As expected, the outcomes of policies targeting all ages lie between the results for the youngest and oldest age groups.

[30] This is in line with Conesa, Kitao, and Krueger (2009), which shows that in the absence of age-dependent tax codes, capital income tax simulates age-dependent tax.

Table 3: Impact of Subsidy Policy Experiments

	25-34	35-44	45-54	All ages
Panel (A). Subsidizing business entrants only				
	(1a)	(2a)	(3a)	(4a)
Subsidies (relative to median income)	0.1811	0.1984	0.2156	0.1035
Impact on entrep. pop. share, p.p. change				
Overall	4.19	2.39	1.98	3.93
Low innate ability type	0.59	0.61	0.65	0.54
Mid innate ability type	3.14	1.72	1.43	3.34
High innate ability type	12.07	5.49	2.59	9.24
Impact on capital and output, % change				
Capital, entrep. production	40.77	19.62	12.42	37.72
Output, entrep.	42.71	20.65	12.40	38.56
Output per entrep.	-6.58	-3.83	-4.14	-6.02
Panel (B). Subsidizing all operating business owners				
	(1b)	(2b)	(3b)	(4b)
Subsidies (relative to median income)	0.1121	0.1259	0.1380	0.0500
Impact on entrep. pop. share, p.p. change				
Overall	4.65	2.94	2.46	4.15
Low innate ability type	0.45	0.50	0.53	0.43
Mid innate ability type	3.97	2.50	2.15	3.85
High innate ability type	11.44	5.46	2.08	7.96
Impact on capital and output, % change				
Capital, entrep. production	48.25	27.22	19.13	42.18
Output, entrep.	49.37	27.68	18.43	41.67
Output per entrep.	-6.83	-4.00	-3.99	-4.83

Note: (i) The level of subsidies is normalized by the median income of all the households of all age groups in the benchmark economy. (ii) Impacts for each experiment case are evaluated as the change in either % or p.p. on the corresponding moments relative to the benchmark economy.

Column (4a) is closer to column (1a), suggesting that the overall effect is primarily driven by increased entrepreneurship at younger ages.

Moreover, although subsidizing entrants significantly increases experimentation with entrepreneurship, it does not severely compromise the quality of entrepreneurs. Output per entrepreneur declines by only around 6%, although aggregate entrepreneur population share increases by 3.9 percent—an over 40% increase. With a subsidy, the selection cutoff for becoming an entrepreneur decreases in all dimensions, including both assets and beliefs. However, the subsidy also allows agents to experiment with entrepreneurship for longer periods, reducing the probability that an innately productive agent will quit upon receiving a low signal realization simply as a result of bad luck, which counteracts the drop in average productivity.

In Panel (B), we assume that the government subsidizes all operating entrepreneurs. Naturally, targeting operating entrepreneurs rather than just entrants, who comprise only a subset of operating entrepreneurs, results in a lower subsidy level. We find that outcomes resemble those in Panel (A); however, policy responses tend to be larger when targeting all operating entrepreneurs. This finding suggests that uncertainty persists beyond the initial learning period, maintaining the policy's insurance value.

Our findings indicate that policies aimed at boosting entrepreneurship should prioritize younger individuals and provide insurance for extended periods. One example could be subsidizing young entrepreneurs for a few periods following their entry until uncertainty is largely resolved.

5.2 Flat Business Income Tax Reform

In this section, we modify the business income tax schedule by applying a constant flat rate to all private business income, as illustrated in Figure 10. We adjust the flat rate levels until we identify one that generates government revenue comparable to the benchmark. We discover that a 20% flat business tax rate maximizes revenue among flat rate schedules and is approximately revenue-neutral to the benchmark. Our result—that we hardly find a flat rate that dominates the current progressive tax system—challenges the conventional view that a flat tax reform may be revenue improving since it favors high-productivity entrepreneurs.

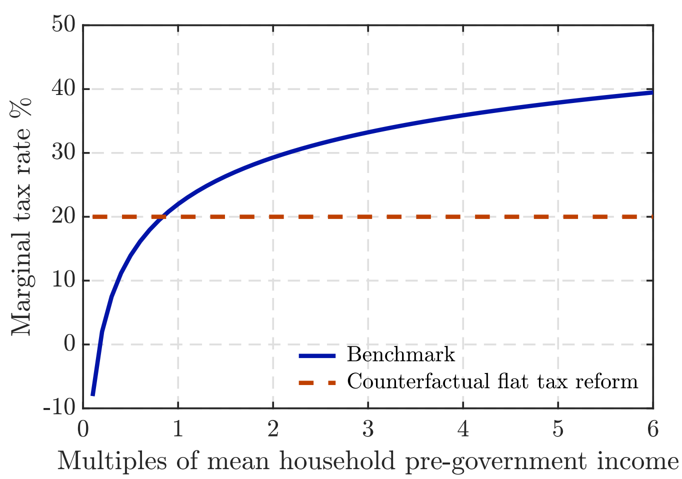

Figure 10: Switch to Flat Business Tax Reform from Benchmark Progressive Tax

Dynamic Effects over the Life Cycle We present the life cycle results in Panel (A) of Table 4 and Figure 11(a). The decline of entrepreneurs in the 25-34 age group is expected, as they face a 29.5% increase in the average tax rate (ATR). Interestingly, older age groups experience an even greater decline in entrepreneurial share, despite lower tax burdens. This is primarily due to the dynamic, persistent effect of learning; fewer individuals discovering their innate productivity in youth results in fewer entrepreneurs in older age groups, as the value of learning decreases with age (Figure 8). Furthermore, entrepreneurial output shifts towards older, wealthier individuals, and the average firm size in the entrepreneurial sector grows.

Table 4: Dynamic Life Cycle Effects of Flat Tax Reform

Age	Aggregate Entre. Share	ATR	Assets	Output
Panel (A). Benchmark with informational frictions and learning				
25-34	–33.6	29.4	5.0	4.7
35-44	–35.7	–1.7	14.2	10.4
45-54	–35.0	–9.0	17.9	11.3
55-64	–38.0	–16.0	26.0	16.4
65-74	–43.0	–20.0	36.9	22.6
Panel (B). Perfect information				
25-34	–36.6	36.3	14.0	12.8
35-44	–19.8	3.7	7.2	7.5
45-54	–14.8	-6.0	10.7	8.8
55-64	–13.3	-11.5	15.5	11.0
65-74	–8.5	-14.7	14.2	8.5

Note: The numbers mean % change in ATRs and entrepreneurial activities (population share, assets, and output) for each age group when tax schedule changes from the benchmark estimation to the counterfactual 20% flat rate.

Distributional Effects across Innate Ability Types Next, we examine the distributional effects of flat business income tax reform across innate entrepreneurial ability types, considering seven productivity levels ranging from -3 to +3 standard deviations from the mean, as in the previous section. We evaluate the percentage point change in entrepreneur share and the percentage change in lifetime total incomes (the sum of wage, business income, and asset income) for each innate productivity level. As shown in Panel (A) of Table 5, agents with highest innate entrepreneurial ability lose most from the flat tax reform. This occurs because the flat tax discourages young agents from exploring their entrepreneurial talents, negatively affecting those with high innate productivity, while

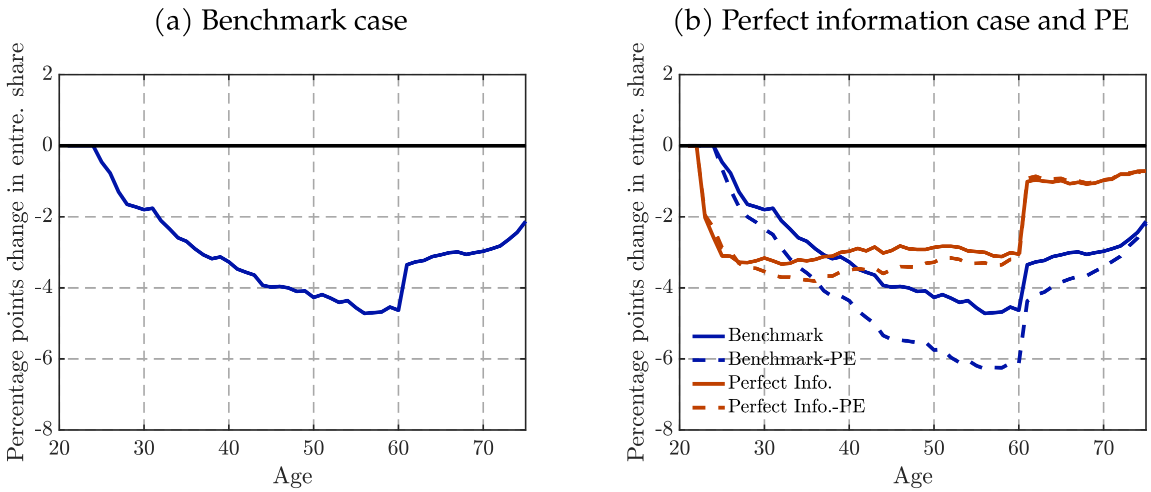

Figure 11: Dynamic Life Cycle Effects of Flat Tax Reform on Entrepreneur Share

Note: Panel (a) plots percentage point change in entrepreneur population share with respect to age when tax schedule changes from the benchmark estimation to the counterfactual 20% flat rate. Panel (b) compares the results in Panel (a) with the cases of perfect information (red curves) and the partial equilibrium (dashed curves).

low-ability agents are unlikely to become entrepreneurs, regardless of learning.

This finding contrasts with the implications of occupation choice models with infinite horizons, such as Evans and Jovanovic (1989) and Cagetti and De Nardi (2006), and the conventional view that high-productivity, high-income entrepreneurs should benefit more from revenue-neutral flat tax reform. Indeed, old and successful incumbent entrepreneurs prefer a flat tax. However, if we take the entire life cycle learning dynamics into consideration, without experimenting with entrepreneurship to learn about their innate entrepreneurial ability, those highly productive entrepreneurs, who are more likely to create high-growth "gazelle" businesses, as emphasized in Sterk, Sedláček, and Pugsley (2021), may never show up, and agents with low entrepreneurial productivity would spend little time being an entrepreneur anyway regardless of the tax policy change.

Aggregate Effects Overall, under the revenue-neutral flat tax reform, even though the average marginal tax rate drops from 26.0% to 24.1%, the aggregate entrepreneur share decreases from 9.0% to 6.0%. Since individuals with high innate productivity, who would have become entrepreneurs under the benchmark progressive tax scheme, either never pursue entrepreneurship or spend less time producing during their lifetime under the flat tax reform, entrepreneurial production falls, which leads to a 1.6% of the decline in aggregate output. Consequently, overall welfare worsens, with consumption-equivalent welfare declining by 2.0%.

Comparison with the Case of Perfect Information Finally, we compare the impacts of the flat tax reform in our benchmark model to an alternative framework with perfect information, as discussed in Section 4. The results are presented in in Panel (B) of Table 4 and Figure 11(b). In the perfect information case, the dynamic persistent effect is less pronounced; only the share of entrepreneurs in the young age group experience a significant decline, with the magnitude of this decline decreasing in older age groups. This occurs because when agents have perfect knowledge of their innate entrepreneurial ability before entering the labor market, it is less important for them to become entrepreneurs at a younger age, as they do not need to learn about their productivity as in the benchmark model. Furthermore, we find that partial equilibrium strengthens the impacts of the flat tax reform in both cases, as illustrated in the right panel of Figure 11(b).

Table 5: Distributional Effects of Flat Tax Reform (General Equilibrium)

Innate ability types	-3 sd	-2 sd	-1 sd	0 sd	+1 sd	+2 sd	+3 sd
Panel (A). Benchmark with informational frictions and learning							
Lifetime entrepreneur share, p.p.	-0.52	-0.72	-1.18	-2.59	-4.44	-7.15	-7.76
Lifetime incomes, %	-1.15	-1.15	-1.30	-2.11	-3.82	-6.93	-8.00
Panel (B). Perfect information							
Lifetime entrepreneur share, p.p.	0	0	0	-0.91	-4.35	-9.12	-4.30
Lifetime incomes, %	0.01	0.03	0.02	-0.42	-1.55	-2.60	2.60

Regarding distributional effects across innate ability types under the perfect information case, as displayed in the Panel (B) of Table 5, the share of lifetime spent as an entrepreneur is still most reduced for agents with the highest innate entrepreneurial ability. However, these agents also experience the most significant gains in terms of lifetime income due to redistribution effects. The contrasting impacts of flat tax reform with perfect information underscore the importance of incorporating entrepreneurial productivity learning dynamics under incomplete information into models used to evaluate tax policies related to entrepreneurship.

6 Conclusions

In this paper, we study the role of uncertainty and learning in selection into entrepreneurship at various life stages and the corresponding macroeconomic and policy implications. Our model is disciplined by both direct and indirect evidence of learning and experimentation with entrepreneurship. We show that incorporating life cycle learning dynamics into the model is crucial for rationalizing the entry and exit dynamics of entrepreneurship by age. More importantly, it changes our perspective on policies aimed at promoting entrepreneurship. Our quantitative results highlight the importance of age in entrepreneurial entry—without early experimentation, highly successful entrepreneurs may never emerge.

The key findings of our paper have broad implications and present directions for future research and policy design. Empirically, the dynamic effect over the life cycle complicates the identification of the causal relationship between tax progressivity and entrepreneurship across time, since reducing tax progressivity to boost current entrepreneurs may come at the expense of future generations' entrepreneurs.

Theoretically, our framework can potentially contribute to the ongoing debate on the sources of the secular decline in US entrepreneurship over the past three decades. Various sources suggest different policy implications. One strand of the literature highlighting the role of skill-biased technological change (e.g., Salgado (2020) and Jiang and Sohail (2023)) argues that the decline in entrepreneurship is an efficient consequence of technological improvement and need not concern policymakers. However, recent empirical and quantitative studies, such as Haltiwanger, Jarmin, and Miranda (2012), Decker et al. (2016), Alon et al. (2018), Sedláček and Sterk (2019), and Karahan, Pugsley, and Şahin (2022), emphasize that the decline is most prominent among the young. Viewed through the lens of our model, if this decline is primarily driven by a large-scale policy change, such as a decrease in insurance value provided by tax and transfer systems, or a macro shock disproportionately hurting the young, it should receive greater attention from the government and policymakers, as such a decline will be propagated by the persistent dynamic effect of learning highlighted in our paper. We leave a more thorough and rigorous analysis of these issues for future research.

References

Alon, Titan, David Berger, Robert Dent, and Benjamin Pugsley. 2018. "Older and Slower: The Startup Deficit's Lasting Effects on Aggregate Productivity Growth." *Journal of Monetary Economics* 93:68–85.

Altig, David, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Brent Meyer, and Nicholas Parker. 2022. "Surveying Business Uncertainty." *Journal of Econometrics* 231 (1):282–303.

Arkolakis, Costas, Theodore Papageorgiou, and Olga A. Timoshenko. 2018. "Firm Learning and Growth." *Review of Economic Dynamics* 27:146–168.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship." *American Economic Review: Insights* 2 (1):65–82.

Bell, Felicitie C. and Michael L. Miller. 2005. *Life Tables for the United States Social Security Area, 1900-2100*. 120. Social Security Administration, Office of the Chief Actuary.

Benabou, Roland. 2002. "Tax and Education Policy in a Heterogeneous-agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2):481–517.

Bhandari, Anmol, Tobey Kass, Thomas J. May, Ellen R. McGrattan, and Evan Schulz. 2023. "On the Nature of Entrepreneurship." Working paper.

Bhandari, Anmol and Ellen R. McGrattan. 2021. "Sweat Equity in US Private Business." *Quarterly Journal of Economics* 136 (2):727–781.

Boar, Corina, Denis Gorea, and Virgiliu Midrigan. 2022. "Why Are Returns to Private Business Wealth So Dispersed?" NBER Working Paper.

Boar, Corina and Matthew Knowles. 2020. "Optimal Taxation of Risky Entrepreneurial Capital." Working Paper.

Boar, Corina and Virgiliu Midrigan. 2023. "Should We Tax Capital Income or Wealth?" *American Economic Review: Insights* 5 (2):259–274.

Boháček, Radim and Jozef Zubrickỳ. 2012. "A Flat Tax Reform in an Economy with Occupational Choice and Financial Frictions." *Economic Journal* 122 (565):1313–1345.

Brüggemann, Bettina. 2021. "Higher Taxes at the Top: The Role of Entrepreneurs." *American Economic Journal: Macroeconomics* 13 (3):1–36.

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

Cagetti, Marco. 2003. "Wealth Accumulation over the Life Cycle and Precautionary Savings." *Journal of Business & Economic Statistics* 21 (3):339–353.

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

Chen, Cheng, Tatsuro Senga, Chang Sun, and Hongyong Zhang. 2023. "Uncertainty, Imperfect Information, and Expectation Formation over the Firm's Life Cycle." *Journal of Monetary Economics* .

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review* 99 (1):25–48.

Cullen, Julie Berry and Roger H. Gordon. 2002. "Taxes and Entrepreneurial Activity: Theory and Evidence for the U.S." NBER Working Paper No. 9015.

David, Joel, Hugo Hopenhayn, and Venky Venkateswaran. 2016. "Information, Misallocation and Aggregate Productivity." *Quarterly Journal of Economics* 131 (2):943–1005.

De Nardi, Mariacristina. 2004. "Wealth Inequality and Intergenerational Links." *Review of Economic Studies* 71 (3):743–768.

DeBacker, Jason, Vasia Panousi, and Shanthi Ramnath. 2023. "A Risky Venture: Income Dynamics among Pass-Through Business Owners." *American Economic Journal: Macroeconomics* 15 (1):444–474.

Decker, Ryan A., John C. Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2016. "Where Has All The Skewness Gone? The Decline in High-Growth (Young) Firms in The US." *European Economic Review* 86:4–23.

Dillon, Eleanor W. and Christopher T. Stanton. 2018. "Self-employment Dynamics and the Returns to Entrepreneurship." Working paper.

Dyrda, Sebastian and Benjamin Pugsley. 2020. "Taxes, Private Equity, and Evolution of Income Inequality in the US." Working paper.

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

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

Feenberg, Daniel and Elisabeth Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1):189–194.

Fossen, Frank M. 2009. "Would a Flat-rate Tax Stimulate Entrepreneurship in Germany? A Behavioural Microsimulation Analysis Allowing for Risk." *Fiscal Studies* 30 (2):179–218.

Gentry, William M. and R. Glenn Hubbard. 2004. "Entrepreneurship and Household Saving." *Advances in Economic Analysis and Policy* 4 (1):1–55.

Guvenen, Fatih. 2007. "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review* 97 (3).

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, and Sergio Ocampo. 2023a. "Taxing Wealth and Capital Income when Returns are Heterogeneous." Working paper.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen. 2023b. "Use it or Lose It: Efficiency and Redistributional Effects of Wealth Taxation." *Quarterly Journal of Economics* 138 (2):835–894.

Guvenen, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer. 2020. "Multidimensional Skill Mismatch." *American Economic Journal: Macroeconomics* 12:210–214.

Haltiwanger, John, Ron Jarmin, and Javier Miranda. 2012. "Where Have All the Young Firms Gone?" *Business Dynamics Statistics Briefing* 6.

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

Hamilton, Barton H., Andrés Hincapié, and Mostafa Salari. 2022. "Entrepreneurial Dynamics and the Small-Firm Effect." Working Paper.

Hamilton, Barton H., Nicholas W. Papageorge, and Nidhi Pande. 2019. "The Right Stuff? Personality and Entrepreneurship." *Quantitative Economics* 10 (2):643–691.

Hansen, Gary D. 1993. "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked." *American Economic Review* 8 (1):71–80.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante. 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967–2006." *Review of Economic Dynamics* 13 (1):15–51.

Heathcote, Jonathan, Kjetil Storesletten, and Gianluca Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *Quarterly Journal of Economics* 132 (4):1693–1754.

Hincapié, Andrés. 2020. "Entrepreneurship over the Life Cycle: Where are the Young Entrepreneurs?" *International Economic Review* 61 (2):617–681.

Hopenhayn, Hugo and Galina Vereshchagina. 2009. "Risk Taking by Entrepreneurs." *American Economic Review* 99 (5):1808–1830.

Humphries, John Eric. 2019. "The Causes and Consequences of Self-Employment over the Life Cycle." Working paper.

Hurst, Erik and Benjamin Pugsley. 2011. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity* :73–143.

———. 2017. "Wealth, Tastes, and Entrepreneurial Choice." *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges* 75:111.

Imrohoroğlu, Ayşe, Cagri S. Kumru, and Arm Nakornthab. 2018. "Revisiting Tax on Top Income." Working paper.

Jiang, Helu and Faisal Sohail. 2023. "Skill-biased Entrepreneurial Decline." *Review of Economic Dynamics* 48:18–44.

Jones, John Bailey and Sangeeta Pratap. 2020. "An Estimated Structural Model of Entrepreneurial Behavior." *American Economic Review* 110 (9):2859–2898.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* :649–670.

Karabarbounis, Marios. 2016. "A Road Map for Efficiently Taxing Heterogeneous Agents." *American Economic Journal: Macroeconomics* 8 (2):182–214.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2022. "Demographic Origins of the Startup Deficit." Working Paper.

Keane, Michael P. and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3):473–522.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives* 28 (3):25–48.

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

Kochen, Federico. 2022. "Finance Over the Life Cycle of Firms." Working Paper.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2020. "The Tail that Wags the Economy: Beliefs and Persistent Stagnation." *Journal of Political Economy* 128 (8):2839–2879.

Lise, Jeremy and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review* 110 (8):2328–2376.

Lockwood, Lee M. 2018. "Incidental Bequests and the Choice to Self-insure Late-life Risks." *American Economic Review* 108 (9):2513–2550.

Manso, Gustavo. 2016. "Experimentation and the Returns to Entrepreneurship." *Review of Financial Studies* 29 (9):2319–2340.

Meh, Césaire A. 2005. "Entrepreneurship, Wealth Inequality, and Taxation." *Review of Economic Dynamics* 8 (3):688–719.

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

Nagypál, Éva. 2007. "Learning by Doing vs. Learning about Match Quality: Can We Tell Them Apart?" *Review of Economic Studies* 74 (2):537–566.

Papageorgiou, Theodore. 2014. "Learning Your Comparative Advantages." *Review of Economic Studies* 81 (3):1263–1295.

Quadrini, Vincenzo. 1999. "The importance of entrepreneurship for wealth concentration and mobility." *Review of Income and Wealth* 45 (1):1–19.

———. 2000. "Entrepreneurship, Saving, and Social Mobility." *Review of Economic Dynamics* 3 (1):1–40.

Queiró, Francisco. 2022. "Entrepreneurial Human Capital and Firm Dynamics." *Review of Economic Studies* 89 (4):2061–2100.

Robinson, Baxter. 2023. "Risky Business: The Choice of Entrepreneurial Risk under Incomplete Markets." Working paper.

Salgado, Sergio. 2020. "Technical Change and Entrepreneurship." Working paper.

Sedláček, Petr and Vincent Sterk. 2019. "Reviving American Entrepreneurship? Tax Reform and Business Dynamism." *Journal of Monetary Economics* 105:94–108.

Senga, Tatsuro. 2018. "A New Look at Uncertainty Shocks: Imperfect Information and Misallocation." Working Paper.

Sterk, Vincent, Petr Sedláček, and Benjamin Pugsley. 2021. "The Nature of Firm Growth." *American Economic Review* 111 (2):547–579.

Weinzierl, Matthew. 2011. "The Surprising Power of Age-dependent Taxes." *Review of Economic Studies* 78 (4):1490–1518.

Wen, Jean-François and Daniel V Gordon. 2014. "An Empirical Model of Tax Convexity and Self-employment." *Review of Economics and Statistics* 96 (3):471–482.

Online Appendix

The Macroeconomic Implications of Uncertainty and Learning for Entrepreneurship

Han Gao
UNSW

Lichen Zhang
HKU

A Data and Measurement

A.1 PSED

We use PSED-I (1998-2004) where there are 590 Nascent Entrepreneurs (NE) and 227 people in Controlled Group (CG) are surveyed. Variables related to businesses include business status, capital structure, legal form, expectations, and performances (sales/employment).

To be considered as a NE, individuals need to satisfy the following four criteria. First, the individual had to currently consider himself or herself as involved in the firm creation process. Second, he or she had to have engaged in some business startup activity in the past 12 months. Third, the individual had to expect to own all or part of the new firm being created. Fourth, the initiative, at the time of the initial screening survey, could not have progressed to the point that it could have been considered an operating business.

Key features of NEs in PSED In terms of legal forms, more than 84% are passthroughs. 50% of NE go with Sole Proprietorships, 20% go with Partnerships, 14% go with S-corp or LLC, 11% go with C-corp, 5% undecided. Regarding whether NEs are attached to paid jobs, about half of them have a paid job (partime or fulltime). 31% of men and 25% of women work full time on their new businesses (>= 35 hrs per week). Large majority of both sexes work for a paid job: Of the 70% of men working for pay, 55% did so full time. The analogous statistics for women are 62% and 39%. In terms of business size operated by NEs, around 40% of men and 50% of women choose to be "merely" self-employed, while the rest expect to become employers over the first five years of operation. As for the industrial choice, a large fraction of the men (35%) is starting a business in Health, Education, and Social services. Among the female NE this is also a strong category (20%). Retail and Restaurants account for 28% of the men and 45% of the women. 15% of the women and 8% of the men chose manufacturing.

Expectation formation and learning The learning process is captured by how forecast revision on a business' performance depends on the corresponding forecast error. We rely on the following questions from PSED I to measure forecast errors and forecast revision. Respondents in Wave 1 of PSED I report (1) *We would like to ask about your expectations regarding the future of this new firm. First, what would you expect the total sales, revenues, or fees to be in the first full year of operation?* (2) *And what about in the FIFTH year?*. Respondents in Wave 2-4 report (1) *What sales or revenue do you expect in the (current financial year/first full year of operation)?* [1] (2) *What annual sales or income would you expect for the firm FIVE years after the first full year of sales?* (3) *What annual sales or income would you expect for the firm TEN years after the first full year of sales?* This means we have (1) forecast data on sales for period 1 and 5 in period 0, (2) realized sales in period 1, 2, 3, and (3) forecast data on sales for period 5 and 10 in period 1,2,3.

(a) Forecasts (b) Realizations

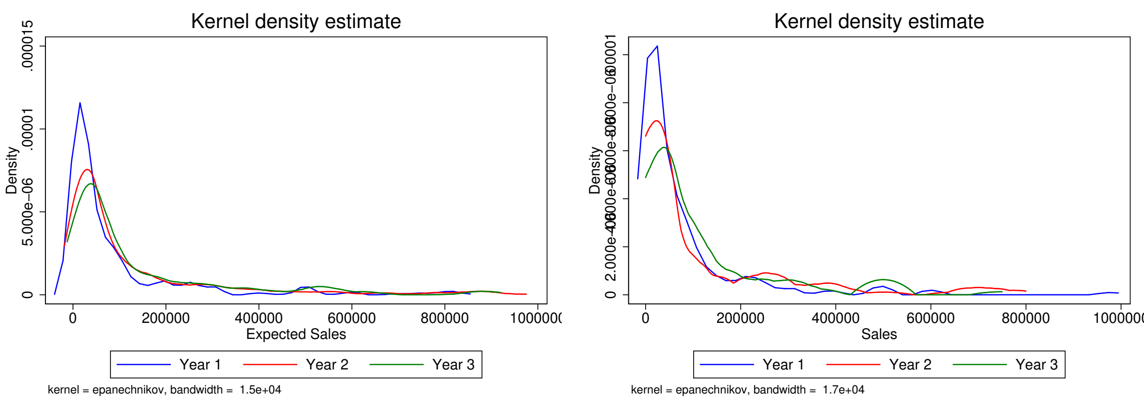

Figure A1: Expected and Realized Sales by Wave

(a) Sales (b) Employment (c) Correlation

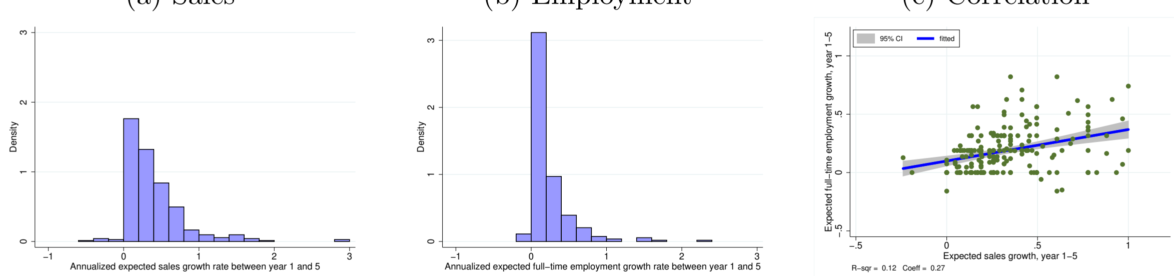

Figure A2: Distribution of Forecast Variables in PSED

[1]Since this question asks about sales/revenue in the current financial year, we regard the answer to this question as the realized sales in that year. To make this approximation solid, we only keep NE that have started operating businesses.

Figure A1 plot the forecasts and realizations of sales across three waves. As Figure A1 shows, both the distributions of forecasts and realized sales are quite dispersed, with long right tails. Besides sales, the PSED also asks forecasts of employment. Figure A2 plots the distribution of forecasts on sales and employment growth, as well as their correlation, respectively. As can be seen from Figure A2, sales and employment growth are positively correlated. However, the expected employment can only take integer values and we also observe a large fraction of entrepreneurs report a constant expected employment. In the data, approximately half of the NEs opt for "mere" self-employment, meaning they have no employees and no intention to expand their employment size. Therefore, we use sales as the primary object to measure entrepreneurs' belief and its evolution. In Table A1, we further presents key summary statistics of forecasts and realized sales.

Table A1: Summary Statistics of Sales in PSED

	Mean	25%	Median	75%	Max	Std. Dev.	Skewness	Frac. zero sales	Exit rate
Expected sales in wave 1 (\$1000), conditional on entry									
Year 1	214	10	30	100	10,000	823	9.22	0.03	
Year 5	1,789	10	100	350	80,000	7,401	7.40	0.01	
Realized sales in following-up waves (\$1000)									
Wave2	241	5	25	90	10,000	1,004	7.34	0.04	0.50
Wave3	508	10	25	185	25,000	2,817	8.38	0.03	0.16
Wave4	887	11	50	200	45,000	5,502	7.87	0.06	–

In Table A2, we compare the mean and standard deviation of forecast errors (FEs) by several key dimensions including gender, age, education, industry, and entrepreneur's past experience in terms of previous entrepreneurial activities, experience in the current industry, and experience in managerial occupations. We find that on average, female entrepreneurs and entrepreneurs with experience in the current industry have lower FE dispersion, while the difference in terms of FE dispersion is not significant in all other dimensions.

Figure 3(b) in the main text shows that FEs predict forecast revisions (FRs), captured by a simple linear regression with a slope of 0.67 for the whole sample. In Table A3, we extend the analysis by adding more control variables to the regression to see how the *learning speed* may potentially differ by entrepreneurs' characteristics. To be more specific, we add interactions of FEs with gender, education, retail industry, age polynomials, and the respondents' previous entrepreneurial experience as well as experience in the current industry. We find that operating in the retail industry is the only variable that significantly affects the learning speed, while all other variables have little impact, which rationalizes

Table A2: Distributions of Forecasting Errors by Characteristics of Entrepreneurs

Variable		Frac. obs.	Mean	Std. Dev.	Frac. missing	t-test	sd-test
Overall			-0.02	0.48			
Female	Yes	0.45	-0.03	0.44	0.01	0.76	0.13
	No	0.55	-0.01	0.52			
Age>=40	Yes	0.55	-0.04	0.50	0.05	0.54	0.58
	No	0.45	0.01	0.47			
College edu.	Yes	0.46	0.00	0.47	0.01	0.60	0.57
	No	0.54	-0.04	0.50			
Retail ind.	Yes	0.23	-0.03	0.48	0.00	0.90	0.99
	No	0.77	-0.02	0.49			
First business	Yes	0.40	0.01	0.43	0.50	0.97	0.66
	No	0.60	0.01	0.46			
Ind. exp.	Yes	0.86	0.03	0.45	0.50	0.50	0.70
	No	0.14	-0.07	0.40			
Manage exp.	Yes	0.18	-0.08	0.46	0.51	0.42	0.85
	No	0.82	0.03	0.45			

our choice of abstracting from detailed characteristics for both the empirics and the model in our main text. In particular, the fact that the initial distribution of forecast errors changes little with respect to entrepreneurs' age indicates that these errors likely pertain to uncertainty about entrepreneurial ability rather than other factors such as demand. This evidence further supports our "no cross-learning between occupations" assumption discussed in Section 2 of the main text.

Summarize survey questions on personality traits We use Principal Component Analysis to summarize the original 25 questions into several key traits. The construction follows Lise and Postel-Vinay (2020), which summarize multiple questions on detailed skills into three main skills. Consider number n types of main traits, the construction method is as follows:

1. Run PCA on PSED questions and keep the first n principal components;
2. Recover traits indices by recombining predicted principal components in such a way that they satisfy n certain exclusion restrictions;
3. Rescale the constructed traits to lie in [0,1].[2]

[2]Technical details are referred to the orginal paper of Lise and Postel-Vinay (2020).

Table A3: FR to FE Regression by Controls

	(1)	(2)	(3)	(4)
FE	0.606***	0.576*	0.642***	0.724*
	(5.19)	(1.89)	(3.01)	(1.88)
FE× female	0.00531	0.0215	0.162	0.353
	(0.04)	(0.14)	(0.67)	(1.37)
FE× college	0.00546	-0.00310	-0.243	-0.222
	(0.04)	(-0.02)	(-1.01)	(-0.90)
FE× retail	0.293*	0.310*	0.359	0.183
	(1.75)	(1.76)	(0.97)	(0.49)
FE× age		0.000339		
		(0.05)		
FE× age^2		-2.86e-08		
		(-0.04)		
FE× first business			-0.308	
			(-1.23)	
FE× log exp				-0.109
				(-0.86)
Personality controls	No	Yes	No	No
Obs.	146	146	72	61
adj. R^2	0.395	0.373	0.254	0.269

Note: t-statistics are reported in parentheses.

Besides the five traits considered by Hamilton, Papageorge, and Pande (2019), we additionally consider a general trait for running a business.[3] This is to isolate the preference solely for doing business, which is orthogonal to the general OCEAN traits such as risk-taking, social activities, etc.. The restrictions we consider are the questions in column 'Restriction' of Table A4 only reflect the corresponding traits.

Correlations between traits Table A5 reports the correlation between the constructed personality traits.

Table A6 describes differences in the constructed personality traits between men and women and between different age groups. Men have significantly higher scores in openness traits and lower scores in extraversion. Older individuals have significantly higher scores in conscientious trait. All other comparisons are not statistically significant. In particular, the love of business trait does not differ by gender or age.

Note that we are not the first to use the data from PSED to document empirical facts

[3]More details on "OCEAN" can be found on https://en.wikipedia.org/wiki/Big_Five_personality_traits.

Table A4: Correspondence between "OCEAN" and Survey Questions

Personality traits	Description	Restriction
Love of business	general love of business	QL1d
Openness to experience	inventive/curious vs. consistent/cautious	QL1q
Conscientiousness	efficient/organized vs. extravagant/careless	QL1b
Extraversion	outgoing/energetic vs. solitary/reserved	QL1h
Agreeableness	friendly/compassionate vs. critical/rational	QL1x
Neuroticism	sensitive/nervous vs. resilient/confident	QL1i

Table A5: Correlations between Personality Traits

	LoB	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Love of Business	1.0000					
Openness	0.3606	1.0000				
Conscientiousness	0.3237	0.3368	1.0000			
Extraversion	0.1182	0.0056	0.3695	1.0000		
Agreeableness	0.3206	0.5560	0.2088	0.6670	1.0000	
Neuroticism	0.2347	0.6973	0.6665	0.6044	0.8184	1.0000

related to non-pecuniary benefits that determines the entry of entrepreneurship. Hurst and Pugsley (2011) also use PSED and show that the median small business reports starting their business for non-pecuniary reasons. However, their approach is different from ours. They rely on the question *"Why do [or did] you want to start this new business?"*. They took the raw responses to the question and created five broad categories of their own including non-pecuniary reasons and reasons related to the generation of income. The main responses in the non-pecuniary category include "want to be my own boss," "flexibility/set own hours," "work from home," and "enjoy work, have passion for it/ hobby." They find that roughly 50 percent of all respondents reported non-pecuniary benefits as being one of the primary reasons they started their business. We see the results generated using our approach complementary to theirs, and the biggest advantage of our approach is that we can generate a distribution of Love of Business characteristic, which can be used to discipline the non-pecuniary utility in our structural model.

Personality traits and entrepreneurship We plot the distribution of scores of the six personal traits for two groups of individuals in our sample—nascent entrepreneurs and workers (control group). As shown in Figure A3, only the distribution of Love of Business

Table A6: Comparison of Personality Traits by Gender and Age

	By gender			By age		
	Men	Women	p-value	Age < 40	Age ≥ 40	p-value
Love of Business	0.5742	0.5749	0.9538	0.5727	0.5774	0.7189
	(0.0086)	(0.0085)		(0.0088)	(0.0094)	
Openness	0.5016	0.4685	0.0018	0.4823	0.4871	0.6694
	(0.0078)	(0.0072)		(0.0075)	(0.0083)	
Conscientiousness	0.6021	0.6237	0.0410	0.6250	0.6006	0.0311
	(0.0074)	(0.0075)		(0.0083)	(0.0076)	
Extraversion	0.5623	0.6117	0.0000	0.5847	0.5876	0.7984
	(0.0071)	(0.0072)		(0.0078)	(0.0079)	
Agreeableness	0.6203	0.6237	0.7123	0.6174	0.6270	0.3297
	(0.0065)	(0.0066)		(0.0067)	(0.0072)	
Neuroticism	0.5912	0.5946	0.7235	0.5945	0.5908	0.7106
	(0.0063)	(0.0067)		(0.0069)	(0.0071)	
Sample size	379	395		337	337	

scores exhibit significant difference between the two groups of people.

We run Heckman Two-Step Regression to further explore how LoB affects the entrepreneurial choice. In the first stage, we run the following regression:

$$Probit\,(E = 1|Z) = \Phi\,(Z\gamma)$$

where $E = 1$ if the respondent is an entrepreneur and $E = 0$ otherwise. In the second stage, we run the following regression:

$$entrep.\,income = X\beta + u$$

We report the results in Table A7.

A.2 PSID

The PSID sample used for studying the life-cycle behavior of entrepreneurs was generated following Heathcote, Perri, and Violante (2010) in general. From the raw data, we extract a sample of heads of households from the SRC sample based on the waves from 1970 to 1997.

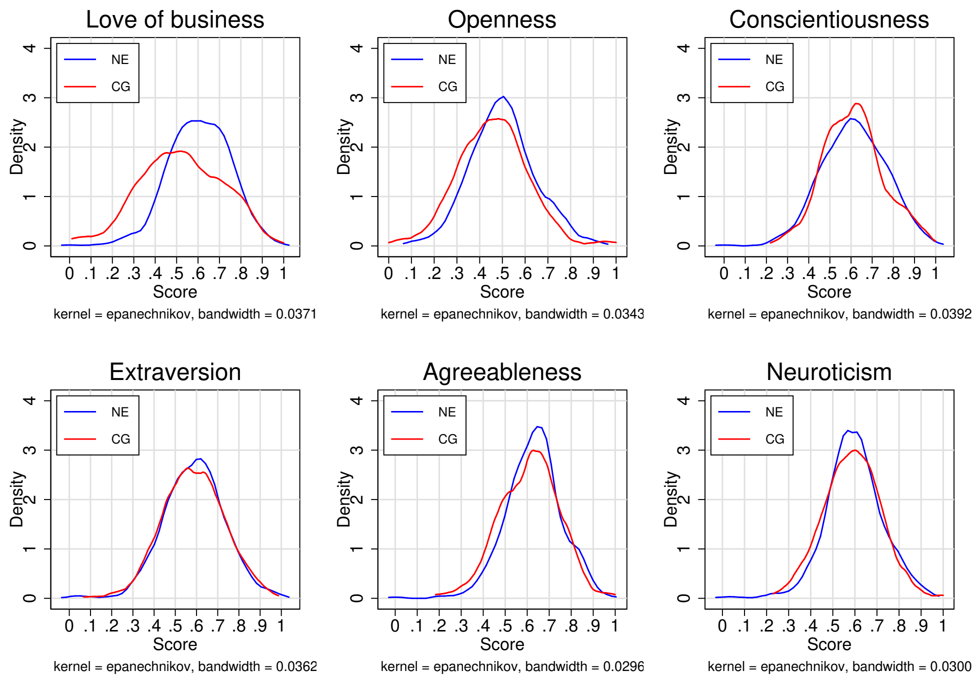

Figure A3: Distribution of Personal Traits Scores

All monetary variables (income and wealth) are deflated using the Personal Consumption Expenditure index (PCE) and expressed in 2010 dollars. The baseline sample considers households whose head is between 21 and 65 years old, both ends included. We report summary statistics of the sample in Table A8.

Definition of "head" The head of the family unit (FU) must be at least 16 years old, and the person with the most financial responsibility in the FU. If this person is female and she has a husband in the FU, then he is designated as head. If she has a boyfriend with whom she has been living for at least one year, then he is head. However, if she has 1) a husband or a boyfriend who is incapacitated and unable to fulfill the functions of head, 2) a boyfriend who has been living in the FU for less than a year, 3) no husband/boyfriend, then the FU will have a female head. A new head is selected if last year's head moved out of the household unit, died or became incapacitated, or if a single female head has gotten married. Also, if the family is a split-off family (hence a new family unit in the sample), then a new head is chosen.

Samples In this paper, we only consider SRC sample (i.e. id68 <= 3000).

Table A7: Heckman Two-Step Regression

		OLS		Heckman Two-step	
Heckman Stage 2 /OLS	Love of Business	0.24	- 0.27	0.21	- 0.49
	Openness	- 1.00	- 0.81	- 0.99	- 0.71
	Conscientiousness	- 0.43	- 1.21	- 0.38	- 0.78
	Extraversion	- 1.68	- 0.78	- 1.70	- 0.76
	Agreeableness	- 0.43	- 2.13	- 0.35	- 1.61
	Neuroticism	2.90	4.50	2.80	3.64
	Age/100	- 0.11	- 0.11	- 0.11	- 0.11
	log(experience)		0.11		0.11
	College	-0.02	- 0.13	- 0.02	- 0.15
	Female	0.05	- 0.13	0.05	- 0.15
	White	0.25**	0.24	0.24	0.21
Heckman Stage 1	Love of Business			2.97***	3.42***
	Openness			- 0.93	- 1.64
	Conscientiousness			- 4.91	- 6.90
	Extraversion			1.24	- 0.32
	Agreeableness			- 7.89	- 8.54
	Neuroticism			10.20	13.95
	Age/100			- 0.02	0.33
	log(experience)				- 0.15**
	White			0.36***	0.52***
	College			0.27**	0.30**
	Female			0.08	0.30*
	Observations	141	70	773	540
	R^2	0.1143	0.1251		

Notes: *,**,*** refer to significance at 10%, 5%, and 1% respectively.

Top-coding and bracketed variables We deal with top-coded observations by assuming the underlying distribution for each component of income is Pareto, and by forecasting the mean value for top-coded observations by extrapolating a Pareto density fitted to the non-top-coded upper end of the observed distribution.

In some of the early waves, a number of income measures were bracketed. For these variables, we use the midpoint of each bracket, and 1.5× the top-coded thresholds for observations in the top bracket.

Variable definitions In the PSID all the questions are retrospective, i.e., variables in survey—year t refer to calendar year $t-1$. The interview is usually conducted around March. When variables were not defined consistently across years (for example employment status was categorized differently in different years), the variables were recoded based on

Table A8: Summary Statistics of PSID sample

	Wage Workers	Entrepreneurs	Labor Force	Total
Obs. per year	2,284	261	2,690	2,994
Age (mean)	38.1	43.0	38.4	39.6
Men (%)	83.4	94.6	83.7	81.4
College or above (%)	27.2	36.1	27.3	26.0
White (%)	89.7	96.0	89.8	89.1
Income (mean, 2010$)	49,357	74,778	50,135	45,882
Wealth (mean, 2010$)	153,164	688,013	206,887	206,294

Notes: The table reports statistics of a sample of heads of households between 21 and 65 years old. Each statistic is the sample average across all the survey waves between 1970 and 1997. Entrepreneurs are defined as self-employed business owners. All monetary values are deflated by the PCE index and expressed in 2010 US dollars.

their original (and less detailed) coding, so as to be consistent across years.

Income and earnings: Labor income of heads is defined as income from wages, salaries, commissions, bonuses, overtime and the labor part of self-employment income. The PSID splits self-employment income into asset and labor components using a 50-50 rule.

The earnings of heads consists of both labor income and business income, which is equal to the labor income of head plus the asset part of business income. Note that the variable on the asset part of business income only applies to individuals who runs unincorporated businesses. Unincorporated business owners are not sheltered from the losses of their ventures through limited liability. This means that a head's income can be positive, zero, or negative.

Wealth: The measure of wealth is the variable WEALTH2, which is available in specific waves of PSID. This variable is constructed as sum of values of several asset types (family farm business, family accounts, assets, stocks, houses, and other real estate etc.) net of debt value.

Annual hours of work is defined as the sum of annual hours worked on the main job, on extra jobs, plus annual hours of overtime. It is computed by the PSID using information on usual hours worked per week and the number of actual weeks worked in the last year.

Labor force: a household head is considered in the labor force if her employment status is either "Working now", "Only temporarily laid off, sick leave or maternity leave", or "Looking for work, unemployed".

Entrepreneur: The PSID provides several questions that can be used to classify individuals' entrepreneurial status. In our analysis, we use two of these questions. The first question is "Did you (or anyone else in the family there) own a business at any time in (year) or have a financial interest in any business enterprise?". The second one is "On your main job, are you (head) self-employed, or are you employed by someone else?". An individual is defined as an entrepreneur if her answer to both questions are "yes".

Worker: a household head is considered to be worker if (1) her employment status is "Working now" or "Only temporarily laid off, sick leave or maternity leave", (2) she is neither self-employed nor a business owner, (3) her labor income is positive, and (4) her annual hours is greater than 260.

Retirement: a household head is considered to be retired if (1) her employment status is "Retired", and (2) her social security income is positive. Note that adding condition (2) is to avoid the misreport of retirement status. If we only rely on condition (1) to define retirement, we will see a pattern that around 5% of retirees in our sample are within the age group of 21-50.

Here is the detailed procedure about constructing variables household earnings and hourly wages:

1. Obtain the SRC sample that includes data for labor income, business income, employment status, gender, age, education, race, wealth, indicator on business owner for heads and wives of households.

2. Drop any observation (household) with missing age for either head or spouse.

3. Drop any observation with missing earnings but positive annual hours of work.

4. Drop any observation with positive earnings but zero annual hours of work.

5. Drop any observation with either head or spouse has nominal wage below half of the minimum wage.

6. Drop any household if neither the head nor the spouse is of working age, which we define as between the ages of 21 and 65.

A.2.1 Earnings and wealth from PSID

In this section, we plot the earnings and wealth over the life-cycle as well as the earnings distributions for different groups of people. We consider three groups: (1) entrants, which means first-year entrepreneurs, (2) incumbents which means entrepreneurs excluding entrants, and (3) workers. Earnings is defined as above. That is, workers' earnings are their labor income. Entrepreneurs' earnings are their labor income plus business income. The measure of wealth is the variable WEALTH2 as found in specific waves of PSID. This variable is constructed as sum of values of several asset types (family farm business, family accounts, assets, stocks, houses, and other real estate etc.) net of debt value.

In Figure A4(a), we plot the median earnings for entrants, incumbents, and workers over the life-cycle respectively, with the median earnings of workers with age 26-30 normalized to one. We can see that the median earnings of entrant entrepreneurs is always smaller than that of workers over the life-cycle. This may suggest non-pecuniary value of entry, which is consistent with the two elements—learning and Love of Business characteristics—in our model.[4]

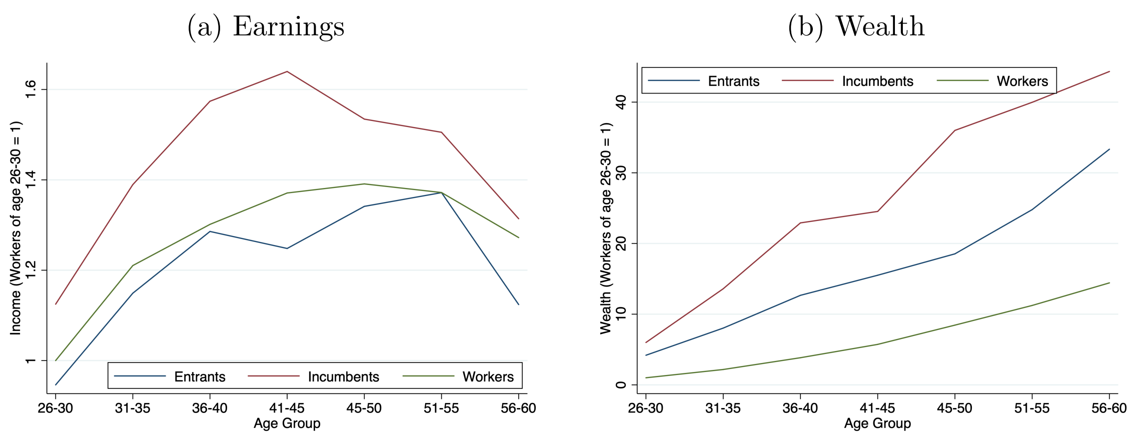

Figure A4: Median Earnings and Wealth over the Life Cycle

In Figure A4(b), we plot the median wealth for entrants, incumbents, and workers over the life-cycle respectively, with the median wealth of workers with age 26-30 normalized to one. It is not surprising to see that entrepreneurs have higher median wealth level compared to workers, which is consistent with the stylized facts that entrepreneurs in general are relatively wealthier people. For example, in the SCF, even though households headed by entrepreneurs make up only 7 to 8 percent of the population, they own nearly one-third of the wealth in the United States.

In Figure A5, we plot the distributions of earnings for entrants, incumbents, and workers respectively. We can see that both the earnings distributions of entrants and incumbents are more dispersed than that of workers, with the median of entrants' earnings smaller.

[4]We also admit that this pattern may be due to some mechanical reasons. For example, if a person enters entrepreneurship in October and is considered as an entrant entrepreneur in that year, then her earnings equals ten months' worker income and two months' entrepreneurial income. However, due to the limitations of PSID, we cannot rule out this kind of possibility.

While median earnings in entrepreneurs are lower than median wage earnings, a subset of entrepreneurs have very high earnings. This may suggest a learning story, as in, for instance, Hincapié (2020) and Dillon and Stanton (2018), that workers seek to maximize expected lifetime earnings may rationally enter entrepreneurship to learn about their entrepreneurial ability, with the option to exit entrepreneurship as uncertainty resolves, even though their realized earnings during entrepreneurship are often low.

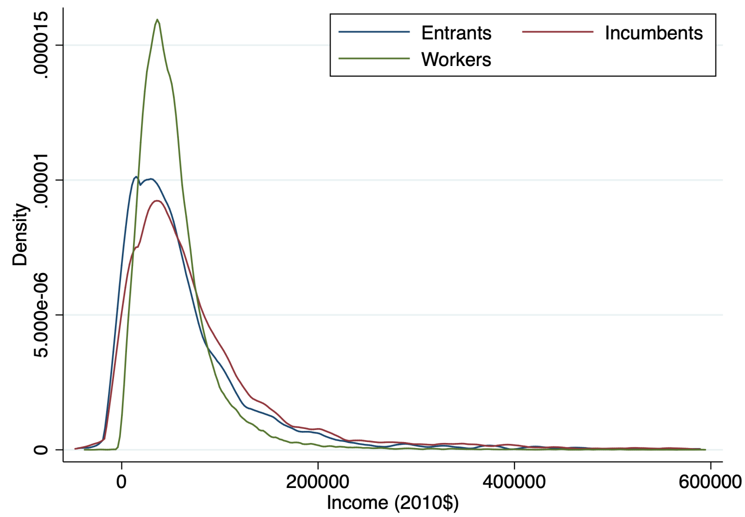

Figure A5: Earnings Distribution in PSID

A.2.2 Life-cycle patterns of entrepreneurship from the CPS

In order to verify the robustness of the life cycle patterns of entrepreneurship, we consider the Current Population Survey (CPS) that covers a much larger number of households compared to PSID. We construct a panel using monthly CPS data following the method developed by Drew, Flood, and Warren (2014). From the raw data, we extract a sample of heads of households from 1976 to 1997 at a monthly basis. All monetary variables (income and wealth) are deflated using the Personal Consumption Expenditure index (PCE) and expressed in 2010 dollars. The baseline sample considers households whose head is between 21 and 65 years old, both ends included. We report summary statistics of the sample in Table A9.

The entry and exit rate can thus only be computed at monthly frequency. We try to make our CPS sample as close to our benchmark PSID sample as possible. The CPS sample covers a similar periods (1975 - 1997) and the entrepreneurs in CPS are defined as self-employed household heads.[5] The age-profiles of entry, exit, and entrepreneurs share

[5]There is no variable on whether an individual is a business owner or not in CPS.

Table A9: Summary Statistics of CPS Sample

	Wage Workers	Entrepreneurs	Labor Force	Total
Obs. per month	32,018	4,777	36,137	42,504
Age (mean)	40.1	44.4	40.5	41.9
Men (%)	75.0	89.9	77.1	73.3
College or above (%)	25.2	29.2	26.0	23.7

Notes: The table reports statistics of a sample of heads of households between 21 and 65 years old. Each statistic is the sample average across all the survey waves between 1976 and 1997 at a monthly basis. Entrepreneurs are defined as self-employed heads of households. All monetary values are deflated by the PCE index and expressed in 2010 US dollars.

using CPS are reported in Figure A6. Although the numbers are not directly comparable between figures using the two datasets due to different definitions of entrepreneurs and different data frequencies, their life-cycle patterns are extremely similar to each other.

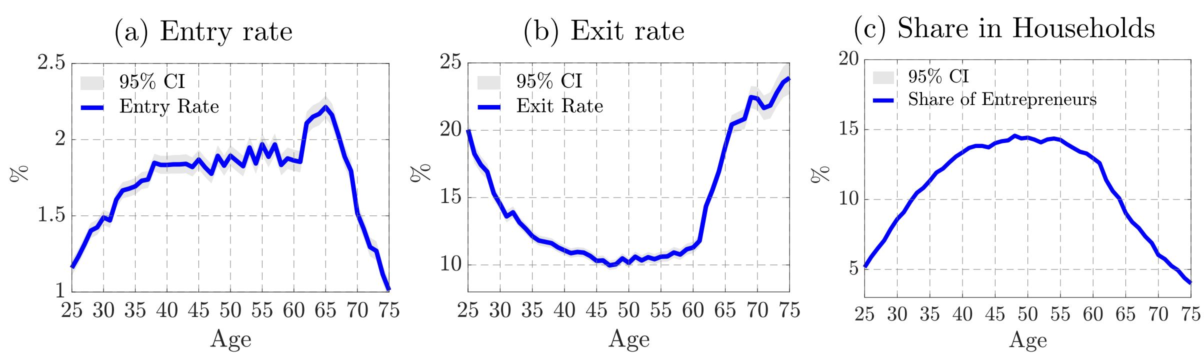

Figure A6: Entrepreneurship over the Life Cycle, CPS Sample

A.3 SCF

We consider two kinds of definitions of business income in SCF and check the share of negative or non-positive business income over the life cycle. The results are reported in Figure A7. In Definition 1, business income = schedule-C business income + taxable interest income + dividend income + capital gains + schedule-E business income + net operating loss. In Definition 2, business income = schedule-C business income + schedule-E business income.

We also document age profiles of the entrepreneur share, earnings, and wealth in SCF, as in Figure A8. The aggregate entrepreneur share over the life cycle is consistent with the patterns in PSID and CPS where we use the same definition of entrepreneurs as the PSID.

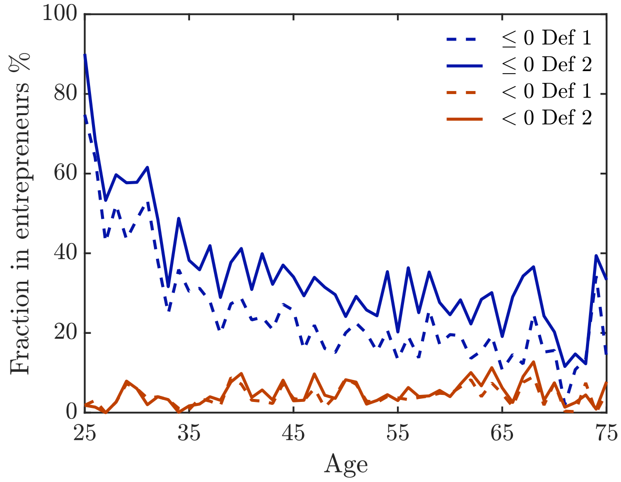

Figure A7: Non-positive Business Incomes in SCF

(a) Entre. share (b) Earnings (c) Wealth

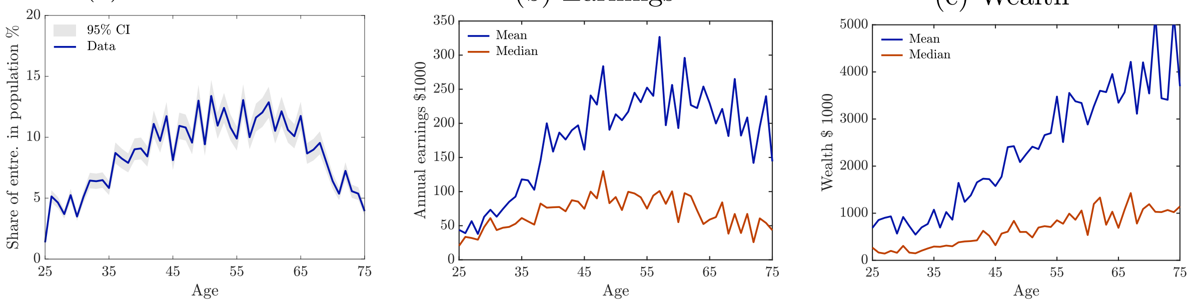

Figure A8: Age Profile of Entrepreneurial Share, Incomes, and Assets in SCF

Finally, we compare several key statistics across PSID, SCF, and PSED as in Table A10. Notes: In IRS integrated business data, share of unincorporated is around 79% in 1996. Among all corporations, around 50% are s-corps.

Table A10: Comparison of Entrepreneurs Sample across PSID, SCF, and PSED

	PSID (96-04)	SCF (97-03)	PSED (98-04)
Frac. of entrep. who have wage income	60%	77%	66%
Frac. of entrep. whose businc>0.5*total inc	49%	56%	-
Share of unincorporated	67%	75%	>70%
Exit rate after 1 year operation	29%	-	30%

B Model

B.1 Details of the Recursive Problems

Value in Normal Working Age $(0 < j < J^V)$ During normal working ages, individuals make occupational choice decisions between being a worker or entrepreneur. For $o = W$,

$$
\begin{aligned}
V_j^W(x_e, \chi_w, a, \epsilon_w, \tilde{\mu}_e, \tilde{\nu}_e, \epsilon_e) &= \max_{a',c,l}\{u(c,l;x_e) \\
&+ \beta[(1-\psi_j)\max_{o'\in\{W,E\}}\{\mathbb{E}V_{j+1}^W(x_e, \chi_w, a', \epsilon_w', \tilde{\mu}_e', \tilde{\nu}_e', \epsilon_e'), \mathbb{E}V_{j+1}^E(x_e, \chi_w, a', \epsilon_w', \tilde{\mu}_e', \tilde{\nu}_e', \epsilon_e')\} + \psi_j\mathcal{V}(a')]\} \\
s.t. \quad & a' + c(1+\tau_c) = a(1+r) + (1-\tau_{ss})y_j^o(a, \epsilon_w, \epsilon_e) - T^o(y_j^o + ra) - \kappa_e \mathbb{1}_{\{o'=E\}} \\
& \{\tilde{\mu}_e', \tilde{\nu}_e'\} = \{\tilde{\mu}_e, \tilde{\nu}_e\} \\
& a' \geq \underline{a},
\end{aligned}
\tag{PA1}
$$

where $y_j^o(a, \epsilon_w, \epsilon_e)$ is the total o-occupation pre-government income, κ_e is the fixed entry cost, and $\mathbb{1}_{\{o'=E\}}$ is the corresponding indicator function that specifies if households who switch occupations to entrepreneurs next period need to pay such a cost.

For $o = E$,

$$
\begin{aligned}
V_j^E(x_e, \chi_w, a, \epsilon_w, \tilde{\mu}_e, \tilde{\nu}_e, \epsilon_e) &= \max_{a',c,l}\{u(c,l;x_e) \\
&+ \beta\delta_e[(1-\psi_j)\mathbb{E}V_{j+1}^W(x_e, \chi_w, a', \epsilon_w', \tilde{\mu}_e', \tilde{\nu}_e', \epsilon_e') + \psi_j\mathcal{V}(a')] \\
&+ \beta(1-\delta_e)[(1-\psi_j)\max_{o'\in\{W,E\}}\{\mathbb{E}V_{j+1}^W(x_e, \chi_w, a', \epsilon_w', \tilde{\mu}_e', \tilde{\nu}_e', \epsilon_e'), \mathbb{E}V_{j+1}^E(x_e, \chi_w, a', \epsilon_w', \tilde{\mu}_e', \tilde{\nu}_e', \epsilon_e')\} + \psi_j\mathcal{V}(a')] \\
s.t. \quad & a' + c(1+\tau_c) = a(1+r) + (1-\tau_{ss})y_j^o(a, \epsilon_w, \epsilon_e) - T^o(y_j^o + ra) \\
& \{\tilde{\mu}_e', \tilde{\nu}_e'\} = \Pi(\tilde{\mu}_e', \tilde{\nu}_e' | \tilde{\mu}_e, \tilde{\nu}_e, \epsilon_e) \\
& a' \geq \underline{a},
\end{aligned}
\tag{PA2}
$$

where δ_e is the exogenous separation shock that only applies to current incumbent entrepreneurs.

B.2 Definition of the Stationary Competitive Equilibrium

An individual with age j is indexed by states $\mathbf{x}_j = (x_e, \chi_w, \ a_j, \ \epsilon_{w,j}, \ \tilde{\mu}_{e,j}, \ \tilde{\nu}_{e,j}, \ \epsilon_{e,j})$. Given a tax structure $\left\{\tau_c, T^\omega(\cdot), T^b(\cdot), \tau_{ss}\right\}$ and an initial distributions of workers and entrepreneurs over individual states $\left\{\Gamma_0^W(\mathbf{x}_0), \Gamma_0^E(\mathbf{x}_0)\right\}$, a **stationary recursive competitive equilibrium** comprises

- prices $\{w, r\}$ and social security benefits z

- a sequence of workers' policy functions on saving, occupation choice, consumption, and hours,
  $\left\{a_W^{'}\left(\mathbf{x}_j\right), o_W^{'}\left(\mathbf{x}_j\right), c_W^{'}\left(\mathbf{x}_j\right), h\left(\mathbf{x}_j\right)\right\}_{j=1}^{J^R-1}$, with associated value functions $\left\{V_j^W\right\}_{j=1}^{J^R-1}$, a sequence of entrepreneurs' policy functions on saving, occupation choice, consumption, capital rental, and labor hired, $\left\{a_E^{'}\left(\mathbf{x}_j\right), o_E^{'}\left(\mathbf{x}_j\right), c_E^{'}\left(\mathbf{x}_j\right), k\left(\mathbf{x}_j\right), n\left(\mathbf{x}_j\right)\right\}_{j=1}^{J^R-1}$, with associated value functions $\left\{V_j^E\right\}_{j=1}^{J^R-1}$, and individuals' policy functions after retirement on saving and consumption, $\left\{a_R^{'}\left(\mathbf{x}_j\right), c_R^{'}\left(\mathbf{x}_j\right)\right\}_{j=J^R}^{J}$, with associated value functions $\left\{V_j^R\right\}_{j=J^R}^{J}$
- factors demand of the corporate sector, $\{K_C, N_C\}$
- a sequence of distributions over idiosyncratic states for both workers and entrepreneurs $\left\{\Gamma_j^W\left(\mathbf{x}_j\right), \Gamma_j^E\left(\mathbf{x}_j\right)\right\}_{j=1}^{J}$

such that

1. Given prices w, r, the tax structure $\left\{\tau_c, T^\omega(\cdot), T^b(\cdot), \tau_{ss}\right\}$, and social security benefits z, the policy functions solve individual's problems (P1) and (P2).
2. The factors demand of the corporate sector solve equation (8).
3. Capital market clears:
$$\sum_{j=1}^{J}\int a^W(\mathbf{x}_j)d\Gamma_j^W(\mathbf{x}_j) + \sum_{j=1}^{J}\int a^E(\mathbf{x}_j)d\Gamma_j^E(\mathbf{x}_j) = K_C + \sum_{j=1}^{J^R-1}\int k(\mathbf{x}_j)d\Gamma_j^E(\mathbf{x}_j) \tag{A1}$$
4. Labor market clears:
$$\sum_{j=1}^{J^R-1}\int \epsilon_{\omega,j}\theta_j h_j(\mathbf{x}_j)\mathbb{I}_{\{h_j>0\}}d\Gamma_j^W(\mathbf{x}_j) = N_C + \sum_{j=1}^{J^R-1}\int n(\mathbf{x}_j)d\Gamma_j^E(\mathbf{x}_j) \tag{A2}$$
5. The Social Security system clears:
$$\tau_{ss}\left(\sum_{j=1}^{J^R-1}\int y_j^\omega(\mathbf{x}_j)d\Gamma_j^W(\mathbf{x}_j) + \sum_{j=1}^{J^R-1}\int y_j^b(\mathbf{x}_j)d\Gamma_j^E(\mathbf{x}_j)\right) = \sum_{j=J^R}^{J} z \tag{A3}$$
6. The government balances its budget:
$$G = \tau_c C + \sum_{j=1}^{J^R-1}\int T^\omega\left(y_j^\omega(\mathbf{x}_j)\right)d\Gamma_j^W(\mathbf{x}_j) + \sum_{j=1}^{J^R-1}\int T^b\left(y_j^b(\mathbf{x}_j)\right)d\Gamma_j^E(\mathbf{x}_j) \tag{A4}$$

7. The distributions of workers and entrepreneurs at the beginning of period j respectively, $\left\{\Gamma_j^W\left(\mathbf{x}_j\right), \Gamma_j^E\left(\mathbf{x}_j\right)\right\}_{j=1}^J$, evolve based on the individuals' policy functions and the autoregressive process for the exogenous productivity states.

B.3 Additional Model Fits of the Benchmark Model

Entrepreneurial Earnings The existing literature that incorporates learning in a structural model of entrepreneurs (e.g., Hincapié (2020); Dillon and Stanton (2018)) typically relies on moments of earnings for entrepreneurs and workers to identify the learning process. We use relatively direct evidence on entrepreneurs' expectation formation to discipline the learning process. In Figure A9, we demonstrate that our model can well replicate the mean and standard deviation of entrepreneurial income with respect to the entrepreneurial spell.

Figure A9: Model Fit: Earnings by Entrepreneurial Duration

Income and Wealth Distributions Our calibrated model captures both the income and wealth distribution well, as in Cagetti and De Nardi (2006). The results are reported in Table A11.

First time entry In Figure A10, we plot the first time entrepreneurs as a share of total population which contains overlapping information with the figure on the entry rate over the life cycle. In addition, we compare the moments implied by the benchmark model with the case of perfect information.

Aggregate moments We further check if the moments on the macroeconomic level generated from our model is consistent with the data. The results are reported in Table A12.

Table A11: Model Fit: Income and Wealth Distribution

	Benchmark	Data
Gini coefficient		
Income – all	0.52	0.55
Income – worker	0.35	0.38
Income – entrepreneur	0.62	0.66
Wealth – all	0.64	0.85
Income/wealth ratios: entrepreneur to worker		
Income – median	1.25	1.30
Income – mean	2.12	2.50
Wealth – median	5.90	6.00
Fraction of entrepreneurs in wealth percentiles		
Top 1%	0.56	0.54
Top 5%	0.48	0.39
Top 10%	0.31	0.32
Top 20%	0.22	0.22

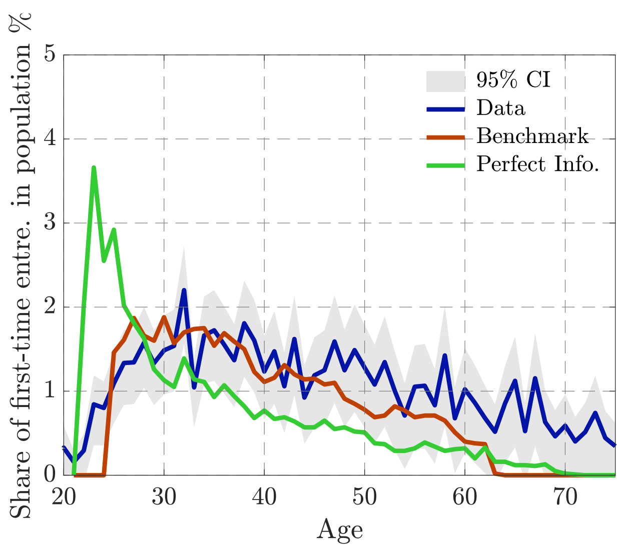

Figure A10: Model Fit: First-time Entrepreneurs as a Share of Total Population

Firm distribution in entrepreneurial sector We finally check the model fitness in distribution of firm size in terms of employment in the entrepreneurial sector. The statistics are reported in Table A13. We can see that our model is able to reproduce similar patterns to the empirical results we obtain from the Survey of Consumer Finance (SCF).

Table A12: Aggregate Moments

	Values
Taxes to GDP ratios, %	
Total taxes	23.9
Consumption tax	2.4
Wage income tax	16.6
Business income tax	1.6
Assets/sales to GDP ratios, %	
Corporate fixed asset	261.6
Entrepreneurial fixed assets	48.3
Entrepreneurial sales	21.3

Table A13: Model Fit: Firm Size Distribution of Entrepreneurs

	Data	Model
Share of entre. in population %	8.8	8.4
Share of hiring entre. %	66.1	82.9
Firm size distribution %		
1-5 Employees	69.2	42.3
6-10 Employees	11.9	40.2
11-20 Employees	6.5	17.5
>20 Employees	12.5	0.0

Exit of entrepreneurs around retirement For the increase in the exit rate after the age of 60, we do a decomposition of the exit rate to distinguish between the exit due to retirement and the exit due to switching occupation. The results are presented in Figure A11. Starting from the age of 62, increasing exit rate is only driven by the increasing retirement of people. Our model well replicates the exit of entrepreneurs around the retirement. The key element that helps to match the data is voluntary retirement and bequest.

Wealth percentiles As shown in Figure A12 using the Survey of Consumer Finance, there is still high level of wealth accumulation at the later stage of individuals' life cycle. By incorporating the element of bequest into our model, the wealth percentiles over the life cycle for all the individuals and for entrepreneurs only are close to its empirical counterpart.

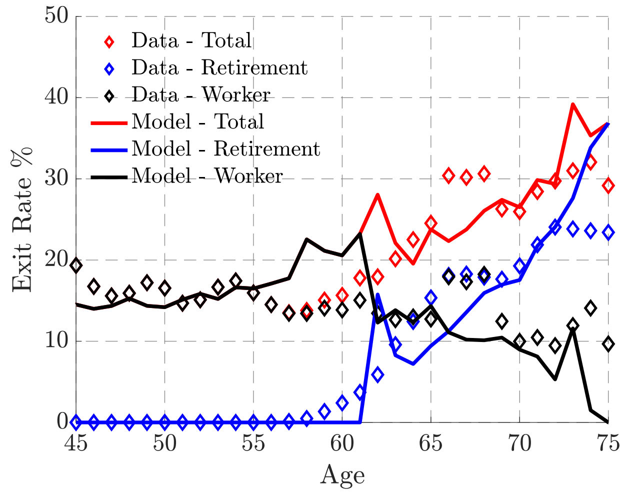

Figure A11: Model Fit: Exit of Entrepreneurs around Retirement

Panel (a) is the data and panel (b) is generated from the model. Our model can replicate the overall life cycle pattern of asset accumulation. To be more specific, in both the data and our model, for the overall population, asset peaks for the age group of 60-64 and drops afterwards. Our model slightly overpredicts the drop in asset for the entrepreneurs at older ages. Our model does a good job in matching the wealth distribution for both the overall population and the entrepreneurs. Our model slightly underestimates the gaps between 95% percentile and the median.

Dispersion of LoB characteristic Results are shown in Table A14.

Table A14: Love of Business Characteristic by Entrepreneur Status

	All	Workers		Entrepreneurs	
		Data	Model	Data	Model
Mean	0.531	0.521	0.524	0.614	0.612
Std. Dev.	0.190	0.193	0.189	0.123	0.171

B.4 Implications of Benchmark vs Perfect Information Model

B.4.1 Re-estimation under the Case of Perfect Information

In this section, we re-calibrate a version of our model, which we refer to as the perfect information case. The only deviation from the benchmark in the perfect information case is that individuals are already aware of their innate entrepreneurial ability upon entering

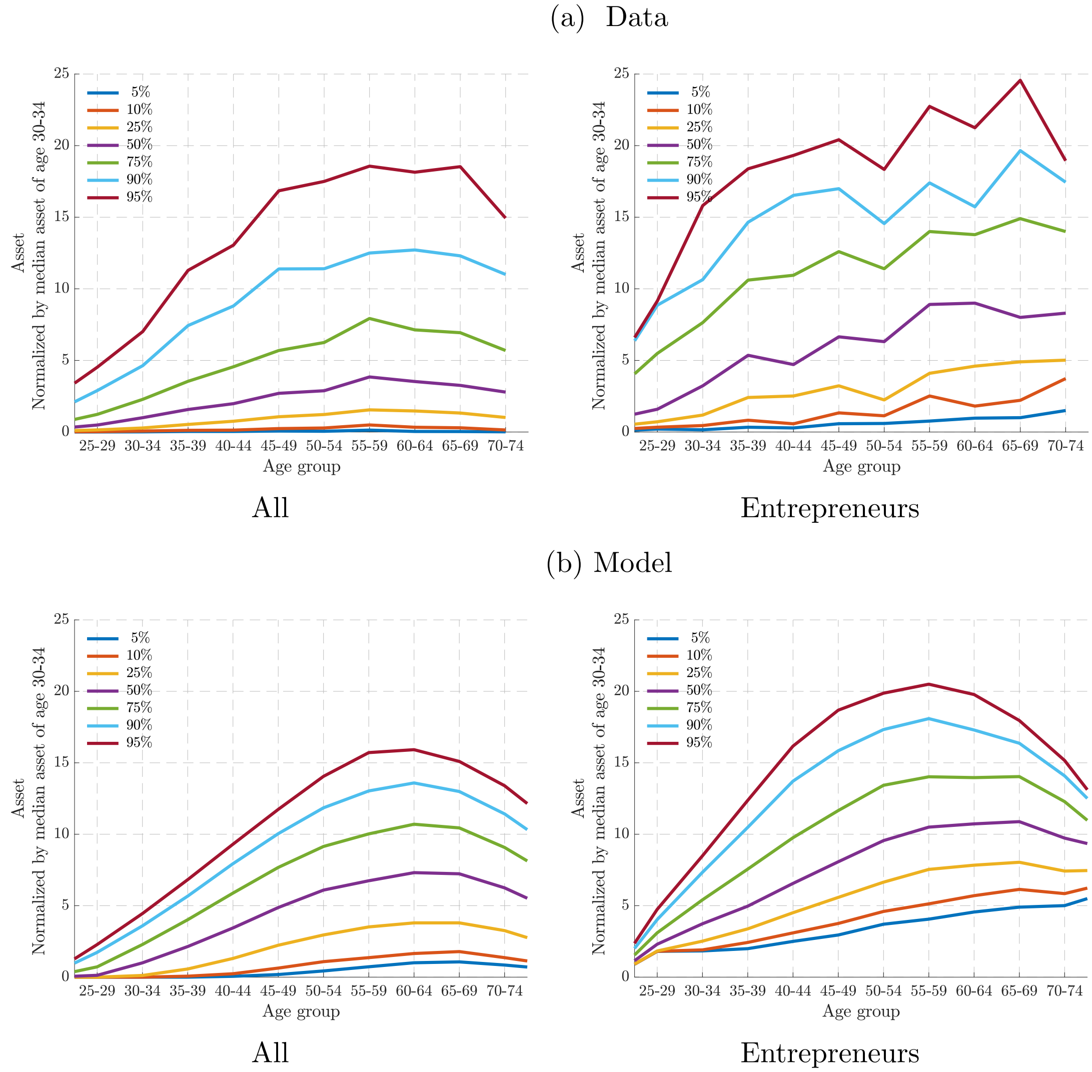

Figure A12: Model Fit: Asset Percentiles over the Life Cycle

the labor market. After they decide to be an entrepreneur, there will still be transitory shocks realized to their innate productivity, which essentially determines the productivity with which output is produced. Thus, our benchmark model can easily nest the perfect information case.

By comparing our baseline estimation results with those of re-estimation under perfect information, we aim to understand how the key elements of our benchmark model—uncertainty arising from imperfect information and learning—affect the parameter estimates on entrepreneurial dynamics. Particularly, we examine how learning contributes to generating a declining exit rate with respect to entrepreneurial spell or working age endogenously.

To proceed, we focus only on the parameters related to entrepreneurial dynamics while fixing the remaining ones to the benchmark estimates. This leaves us with six parameters to re-estimate: $\{\mu_e, \nu_e, \sigma_e, \kappa_o, \kappa_e, \delta_e\}$. We jointly calibrate these parameters to match six empirical moments related to the entrepreneurial sector: (1) population share of entrepreneurs; (2) annual entry rate; (3) annual exit rate; (4) ratio of mean income of entrepreneurs to mean income of workers; (5) ratio of median income of entrepreneurs to median income of workers; and (6) fraction of entrepreneurs with negative incomes.

Table A15: Comparison of Estimates: Benchmark v.s. Perfect Information

Parameter	Description	Benchmark	Perfect Information
μ_e	Mean entrep. productivity	1.31	1.40
ν_e	Std. dev.: innate entrep. prod.	0.17	0.10
σ_e	Std. dev.: i.i.d. shocks	0.29	0.33
κ_o	Per period operational cost	0.02	0.06
κ_e	One-time entry cost	0.05	0.00
δ_e	Exogenous exit rate	0.02	0.09

The main takeaway is that in the perfect information case, where the learning channel driving entrepreneurial choice is absent, entrepreneurial exits are primarily explained by the exogenous separation shock δ_e. This can be observed in the last row of Table A15. Consequently, we do not see the pattern of declining exit rates with respect to working age, especially at young ages.

One additional result we learn from this exercise is that since young agents in our benchmark model are more likely to experiment with entrepreneurship to explore their innate ability, the majority of business losses occur at young ages, and the fraction of entrepreneurs with negative business income decreases by age. In contrast, under the case of perfect information, the fraction slightly increases by age because the entry cutoff decreases as agents accumulate assets over their lifetime.

B.4.2 Entrepreneurship over the Life Cycle by Innate Ability

In Figure A13, we report the share of entrepreneurs in the population with respect to age across various innate entrepreneurial ability types for both benchmark case and perfect information case. In line with our primary findings in the main text, under perfect information, individuals with high innate abilities have a significantly larger share of their lifetime spent as entrepreneurs. This increase is particularly pronounced from a young age.

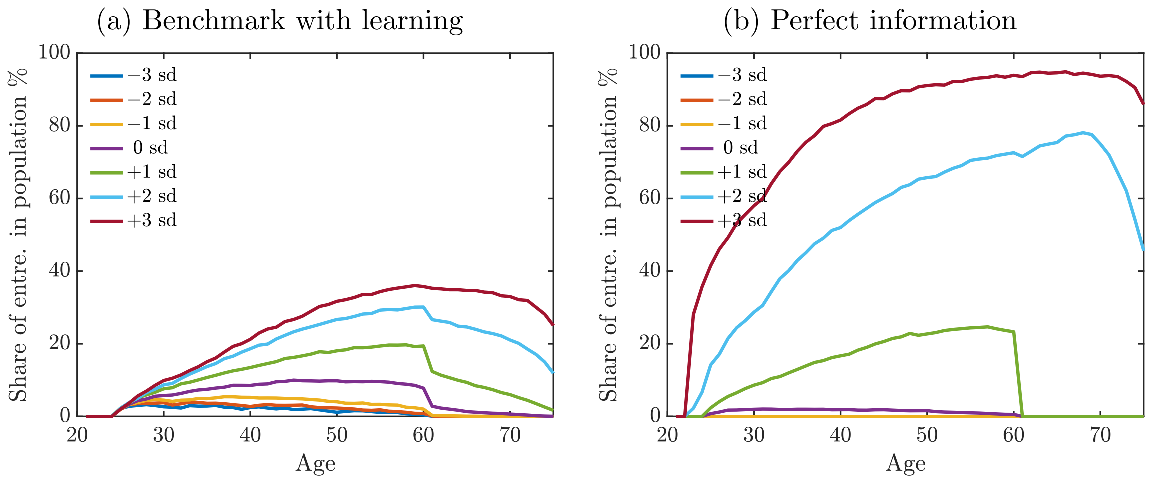

Figure A13: Entrepreneur Share by Innate Productivity Type over the Life Cycle

B.4.3 Interaction between Informational and Financial Frictions

Besides the saving behavior for workers and entrepreneurs in both benchmark and perfect information scenarios reported in the main text, we also investigate the saving behavior of various entrepreneurial ability types in these scenarios. As illustrated in Figure A14, low-ability-type agents display similar hump-shaped saving patterns throughout their life cycle in both cases. In the perfect information scenario, high-ability-type agents save substantially from the very beginning. However, in the benchmark case, due to limited information on their ability types, these agents save less during their early years compared to the perfect information scenario.

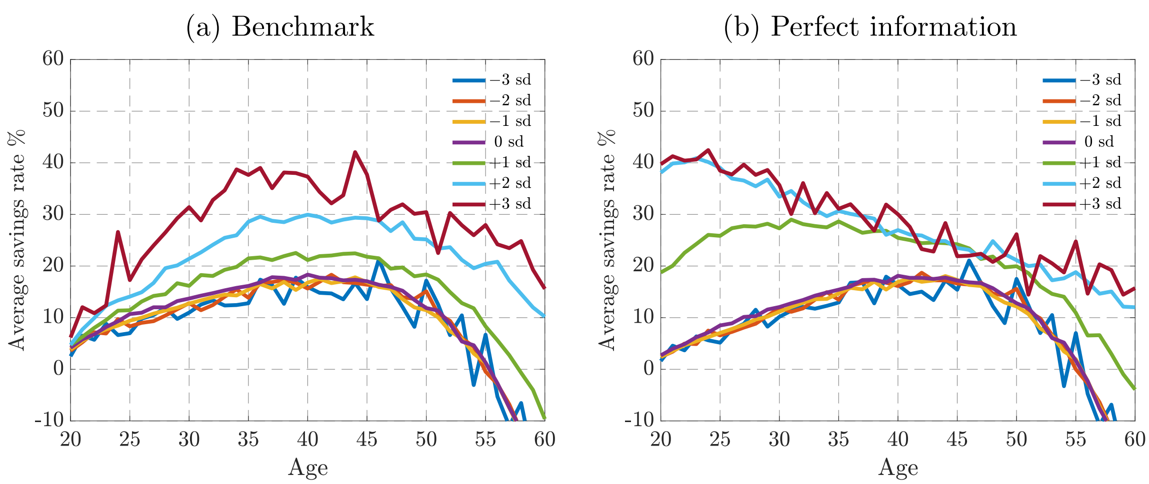

Figure A14: Saving Rates by Innate Entrepreneurial Productivity Types

Next, to help further understand Panel (B) of Figure 9 in the main text, we report the percentage point change in both entry rate and exit rate after the collateral constraint is eased, i.e., λ is raised from 1.5 to 2, in Figure A15.

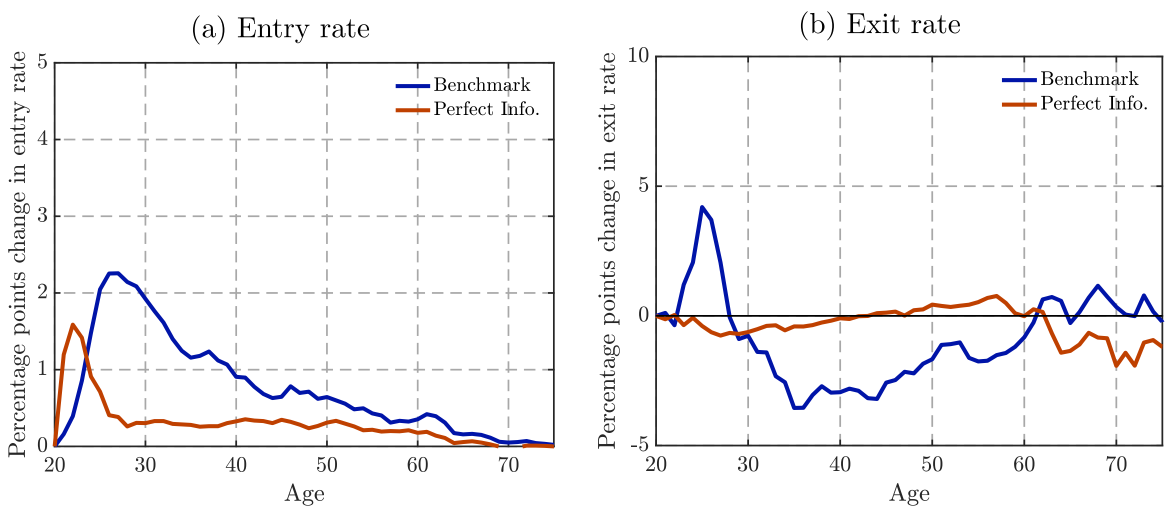

Figure A15: Percentage Point Change before and after the Collateral Constraint is Relaxed

In Table A16, we provide further analysis on the impact of relaxing the collateral constraint on aggregate moments. We report impact on moments when prices and parameters are fixed. Panel A presents data on the life cycle entry and exit of entrepreneurs, as well as the fractions of lending entrepreneurs and entrepreneurs facing constraints. Panel B focuses on aggregate capital and output. As shown in Table A16, under perfect information, where individuals perfectly know their comparative advantage since entering the labor market, the relaxation of the collateral constraint has a much smaller effect on life cycle entrepreneurship, entry and exit rates. Note that, as demonstrated in Table A15, the recalibration will lead to a lower spread of innate ability dispersion, thus the increase in population share of entrepreneurs and average output per entrepreneur would be further attenuated in a recalibrated version.

B.5 Additional Results of Policy Impacts

In Table A17, we present the effects of the revenue-neutral flat business income tax reform on aggregate moments for both the benchmark and perfect information cases. In the benchmark case, which includes informational frictions and learning, the flat tax reform significantly reduces the aggregate share of entrepreneurs and entrepreneurial output, as high innate entrepreneurial ability agents are most affected by this tax policy change. As a result, total output also decreases. However, in the perfect information case, the redistribution effects benefit agents with the highest entrepreneurial ability, despite a

Table A16: Impact on Aggregate Moments When the Collateral Constraint is Relaxed

	Benchmark with learning			Perfect information		
	$\lambda = 1.5$	$\lambda = 2.0$	Δ, p.p.	$\lambda = 1.5$	$\lambda = 2.0$	Δ, p.p.
	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)
Panel A. Entry/exit/collateral-constrained/lending of entrep.						
Entrep. pop. share	0.0969	0.1430	4.61	0.1352	0.1607	2.55
Entry rate	0.0204	0.0306	1.02	0.0165	0.0197	0.32
Exit rate	0.2207	0.2048	-1.59	0.1111	0.1087	-0.24
Frac. of constrained	0.6377	0.5652	-7.25	0.6835	0.5906	-9.29
Frac. of lending	0.7286	0.7249	-0.37	0.7707	0.7507	-2.00
	$\lambda = 1.5$	$\lambda = 2.0$	Δ,%	$\lambda = 1.5$	$\lambda = 2.0$	Δ,%
	(1b)	(2b)	(3b)	(4b)	(5b)	(6b)
Panel B. Aggregate capital and output (normalized)						
Capital, aggregate	1	1.1444	14.44	1.2273	1.3429	9.42
Capital, entre. production	1	1.8901	89.01	1.8942	2.7854	47.04
Capital, lending	1	2.7965	179.65	1.9357	4.2220	118.11
Output, aggregate	1	1.1322	13.22	1.1624	1.2490	7.45
Output, entre.	1	1.7361	73.61	1.9719	2.6473	34.25
Output per entre	1	1.1307	13.07	1.3349	1.4748	10.48

slight decrease in the aggregate share of entrepreneurs. Consequently, the decline in entrepreneurial output is minimal, and total output increases.

Table A17: Impacts of the Revenue-neutral Flat Tax Reform on Aggregate Moments

	Benchmark with learning	Perfect information
Self-employment rate	–36.3%	–16.3%
Interest rate	4.7%	-5.0%
Wage rate	–1.1%	0.9%
Total output	–1.6%	1.8%
Private business	–26.5%	-1.4%
Coporate	16.5%	10.4%
Ave. private business output	16.1%	18.1%
Agg. employee hours	1.0%	1.3%
Agg. capital	5.5%	10.3%
AMTR-worker	1.4%	0.9%
AMTR-entre.	–46.3%	–45.6%
ATR-worker	0.8%	1.0%
ATR-entre.	-15.1%	–12.7%

References

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

Dillon, Eleanor W. and Christopher T. Stanton. 2018. "Self-employment Dynamics and the Returns to Entrepreneurship." Working paper.

Drew, Julia A. Rivera, Sarah Flood, and John Robert Warren. 2014. "Making Full Use of the Longitudinal Design of the Current Population Survey: Methods for Linking Records across 16 months." *Journal of Economic and Social Measurement* 39 (3):121–144.

Hamilton, Barton H., Nicholas W. Papageorge, and Nidhi Pande. 2019. "The Right Stuff? Personality and Entrepreneurship." *Quantitative Economics* 10 (2):643–691.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante. 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967–2006." *Review of Economic Dynamics* 13 (1):15–51.

Hincapié, Andrés. 2020. "Entrepreneurship over the Life Cycle: Where are the Young Entrepreneurs?" *International Economic Review* 61 (2):617–681.

Hurst, Erik and Benjamin Pugsley. 2011. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity* :73–143.

Lise, Jeremy and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review* 110 (8):2328–2376.