Risk-taking over the Life Cycle: Aggregate and Distributive Implications of Entrepreneurial Risk*

Dejanir H. Silva† Robert M. Townsend‡

October 2025

Abstract

We study entrepreneurs' risk-taking over the life cycle in an economy with partial insurance against idiosyncratic shocks. The model replicates the aggregate and idiosyncratic risk premia, the life-cycle patterns of consumption and risk-taking, and the wealth inequality observed in the data. We use the model to evaluate a reform that relaxes risk constraints. The reform lowers the idiosyncratic risk premium and triggers an investment boom. Consistent with a Kuznets-curve dynamic, inequality rises in the short run but falls in the long run. While the initial generation of entrepreneurs benefits from improved insurance, future generations are worse off.

KEYWORDS: Entrepreneurship, risk-taking, risk premium, insurance, inequality

JEL CLASSIFICATION: G11, G51, E44.

*We would like to thank Antonio Antunes, Francisco Buera, Pedro Brinca (discussant), Sylvain Catherine (discussant), David Lagakos, Juan Passadore, and seminar participants at Barcelona GSE Summer Forum, CEPR Household Finance Series, Econometric Society Meeting, FGV-EPGE, LACEA-LAMES, Lubramacro, SED, Tilburg GIFT Conference, UFABC, WFA, and Yale for their helpful comments and suggestions.

†Purdue University, Daniels Business School. Email: dejanir@purdue.edu.

‡Massachusetts Institute of Technology, Department of Economics. Email: rtownsen@mit.edu.

1 Introduction

Entrepreneurship is inherently risky, and entrepreneurial risk is especially important in developing economies, where production is largely carried out in privately owned businesses and risk-sharing opportunities are scarce. Because entrepreneurs hold highly under-diversified portfolios,[1] such risks have far-reaching consequences. At the individual level, they distort investment and savings decisions, with the importance of uninsurable business income varying across entrepreneurs and over the life cycle. These dynamics shape patterns of wealth inequality. At the macro level, limited insurance generates an inefficient risk premium, depressing capital accumulation and hindering economic development. The pervasive effects of incomplete risk sharing underscore the importance of policy interventions that mitigate entrepreneurs' lack of diversification.

This paper studies the aggregate and distributive implications of entrepreneurial risk in a quantitative life-cycle model with imperfect idiosyncratic insurance. We discipline the model using a rich dataset on small businesses in Thailand, which contains detailed information on entrepreneurial returns, portfolio allocations, and consumption.[2] Our model replicates key features of the data, including the risk premium on business returns, the life-cycle profile of risk-taking and consumption, and the inequality patterns both across and within age groups. By combining a theory of the idiosyncratic risk premium with detailed micro data on entrepreneurs' portfolios and returns, we assess the limits to risk sharing and quantify the impact of relaxing these constraints on investment, inequality, and economic development.

We begin by studying empirically the determinants of expected entrepreneurial returns. Since expected returns are closely linked to the marginal product of capital (MPK),

[1] For evidence on under-diversification among entrepreneurs, see, e.g., Moskowitz and Vissing-Jørgensen (2002) and Herranz et al. (2009).

[2] A large literature studies the macroeconomic and asset-pricing implications of firm-level risk (see, e.g., Christiano et al. 2014, Gârleanu et al. 2015, and Herskovic et al. 2018), but these studies typically rely on data from public firms.

the required return on a project is a key determinant of its scale.[3] With imperfect risk sharing, entrepreneurs demand compensation for bearing idiosyncratic risk, whereas under perfect risk sharing expected returns would depend only on aggregate risk exposure. To assess the role of risk-sharing frictions, we estimate a parsimonious two-factor model where expected returns depend on exposures to aggregate and idiosyncratic risk. We find that entrepreneurs more exposed to either source of risk earn higher expected returns. Strikingly, the two-factor model accounts for nearly 70% of the cross-sectional variation in average returns. Since both idiosyncratic volatility and aggregate betas are measured with error, this estimate is a lower bound. Indeed, when entrepreneurs are sorted by aggregate and idiosyncratic risk exposure, nearly 90% of the variation in expected returns can be explained by these two factors. Thus, differences in risk exposure can account for almost all cross-sectional heterogeneity in entrepreneurial returns.

Building on recent advances in empirical asset pricing, we also estimate a richer factor model in which returns are driven by both observable and latent aggregate factors. Expected returns may then reflect compensation for idiosyncratic risk, multiple sources of aggregate risk, and a residual component orthogonal to risk. We find that this residual component plays only a limited role. Thus, a risk-based explanation accounts for nearly all of the variation in expected returns.

While more than 90% of return variance reflects idiosyncratic shocks, less than half of expected returns compensate for such shocks. Sharpe ratios thus differ across sources of risk. This pattern is inconsistent with autarky, where only total variance would matter, and with perfect risk sharing, where idiosyncratic risk would not be priced. The evidence instead suggests that entrepreneurs can only partially insure idiosyncratic shocks.

We then turn to the determinants of entrepreneurs' risk-taking, measured by the share of net worth invested in the business. Risk-taking depends not only on the risk and return of the project but also on attitudes toward risk. Household characteristics shape these

[3]See David et al. (2022) on how differences in aggregate risk premia generate dispersion in MPK.

preferences, and we find substantial life-cycle variation: young entrepreneurs are nearly 40% more exposed to their business than older entrepreneurs. These differences cannot be explained by variation in expected returns, pointing to heterogeneity in risk tolerance across age groups.

To capture these motivating facts, we develop a general equilibrium model with two key ingredients: limited idiosyncratic insurance and finite lives with imperfect altruism. Entrepreneurs operate a technology exposed to aggregate and idiosyncratic shocks. Because of a moral hazard problem, they must bear a fraction of the idiosyncratic risk in equilibrium.[4] We identify the moral hazard parameter by matching the decomposition of expected returns into aggregate and idiosyncratic components in the data. Intuitively, if entrepreneurs could diversify most risks, expected returns would primarily compensate for aggregate risk. Our estimates imply that entrepreneurs can diversify only about 40% of idiosyncratic volatility, leaving substantial exposure that meaningfully shapes investment decisions.

Importantly, we assume that entrepreneurial households also receive labor income from members working outside the business, consistent with the data. As a result, entrepreneurs' effective risk aversion depends inversely on the ratio of human wealth—the present discounted value of future labor income—to financial wealth invested in safe assets and the business. In the data, this human–financial wealth ratio declines over the life cycle. The mechanism endogenously generates heterogeneity in risk aversion, which is essential to match the empirical life-cycle profiles.[5]

Entrepreneurial risk has significant effects on inequality. Wealth inequality rises with age early in life and declines later on. Heterogeneity in returns drives the initial increase, while the rising marginal propensity to consume (MPC) as entrepreneurs near the end of the life cycle reduces dispersion. The model captures this inverted-U pattern in the data.

The model also replicates the heterogeneity in risk premia observed empirically. To

[4] To focus on idiosyncratic risk, we assume only idiosyncratic risk sharing is subject to frictions.

[5] The mechanism is reminiscent of portfolio choice with labor income (see, e.g., Bodie et al. 1992).

generate variation in expected returns across entrepreneurs, we extend the framework along two dimensions: heterogeneous idiosyncratic volatility and decreasing returns to scale (DRS) in production. DRS is essential: without it, the capital–labor ratio would be equalized across entrepreneurs, leaving no systematic differences in expected returns despite variation in risk.

Generating the observed patterns of expected returns with standard ingredients such as skill heterogeneity or collateral constraints proves difficult. A version of the model with heterogeneous skill but no differences in idiosyncratic risk fails to produce dispersion in returns: more productive entrepreneurs operate larger businesses, but expected returns are pinned down by risk exposure. Introducing collateral constraints breaks the link between risk and return, making expected returns depend on net worth for constrained entrepreneurs. Empirically, however, we find no significant relationship between returns and net worth. If anything, risk exposure explains an even larger share of return variation among low-wealth entrepreneurs, who are most likely to be constrained. Taken together, these findings point to a risk-based explanation of entrepreneurial returns.

Having shown that the model captures the main features of the data, we study a counterfactual in which entrepreneurs gain access to better idiosyncratic insurance. This improvement can be interpreted as an expansion of state-contingent credit—for instance, through a flexible lending contract that allows entrepreneurs to delay repayments after negative shocks. Quantitatively, halving the moral hazard parameter lowers the risk premium by 140 basis points and raises the capital stock by about 13%.

Considering the transitional dynamics is essential to evaluate welfare. Despite better diversification, entrepreneurs are worse off in the long run, as lower expected returns slow wealth accumulation. By contrast, the initial generation benefits from the reform: they inherit wealth before the intervention and enjoy a sharp increase in business values. Hence, most of the welfare gains accrue to the initial generation of entrepreneurs and to wage earners, who benefit from higher wages as the capital stock expands.

Related literature. This paper relates to several strands of literature in macroeconomics and finance. First, it connects to work studying how firm-level uncertainty affects asset prices and the real economy (Herskovic et al. 2016, Iachan et al. 2021).[6] While this literature emphasizes business-cycle fluctuations, we instead focus on how firm-level uncertainty shapes long-run outcomes.

Second, it is related to work on how entrepreneurs' lack of diversification affects real investment (Panousi and Papanikolaou 2012), capital structure (Chen et al. 2010, Herranz et al. 2015), and risk-taking (Chen and Strebulaev 2019). This literature is mainly partial equilibrium and abstracts from the aggregate implications that are central to our analysis.

Third, we contribute to research on heterogeneous returns and inequality. This work documents large heterogeneity in portfolio returns (Fagereng et al. 2019, Bach et al. 2020), shows that private businesses are a key source of wealth at the top (Smith et al. 2019, Smith et al. 2020), and highlights the role of return heterogeneity in explaining inequality (Gomez 2017, Hubmer et al. 2021). Our contribution is to show how the idiosyncratic risk premium in private businesses shapes both inequality and aggregate outcomes.

A related strand studies partial insurance of *labor income* (e.g., Kaplan and Violante 2010). Heathcote et al. (2014) show that the degree of partial insurance can be inferred using consumption and labor income data. We show that the degree of partial insurance of entrepreneurial risk can likewise be identified through the idiosyncratic risk premium.

There is also an extensive micro-development literature on risk sharing (Townsend 1994, Morduch 1995) and on the risk and return of production activities in developing economies (Udry and Anagol 2006, De Mel et al. 2008). The macro-development literature, in turn, studies the aggregate implications of credit constraints (Buera and Shin 2013, Midrigan and Xu 2014, Moll 2014). Our approach is complementary, as we emphasize risk constraints rather than borrowing constraints.

Our empirical approach substantially extends that of Samphantharak and Townsend

[6] Another strand studies the asset-pricing implications of labor income risk in infinite-horizon (e.g., Constantinides and Duffie 1996) and life-cycle models (e.g., Storesletten et al. 2007).

(2018). We focus on the risky component of returns—the relevant counterpart to our theory—and show that a parsimonious two-factor model explains most of the cross-sectional variation in returns. We also account for omitted factors and measurement error, document the life-cycle patterns, and develop a structural model linking these empirical findings to partial insurance.

A complementary literature studies how risk shapes entrepreneurial activity in developed economies (e.g., Tan 2018; Robinson 2021; Boar et al. 2022). Peter (2025) shows that equity frictions—a form of imperfect insurance akin to our moral hazard problem—drive entrepreneurial activity and inequality across the Eurozone, indicating that our mechanism is not unique to developing economies.

Our framework builds on Angeletos (2007), extending it to include partial idiosyncratic insurance, a rich demographic structure, and aggregate shocks. These features are essential to match the life-cycle patterns of consumption and risk-taking and to capture the dynamics of inequality after reforms that relax risk constraints.

Organization. The rest of the paper is organized as follows. Section 2 presents empirical evidence on entrepreneurial activity. Section 3 introduces the model. Section 4 analyzes its life-cycle implications. Section 5 studies heterogeneous risk premia. Section 6 examines the effects of changes in idiosyncratic insurance. Section 7 concludes.

2 Motivating evidence

In this section, we provide motivating evidence on entrepreneurial activity in the context of a developing country. First, we study the determinants of entrepreneurial returns. Second, we consider how risk-taking and consumption decisions evolve over the life cycle.

Data. We use data from the Townsend Thai Monthly Survey (TTMS), an intensive monthly survey initiated in 1998 in four provinces of Thailand. Two provinces, Chachoengsao and

Lopburi, are semi-urban in a more developed central region near the capital, Bangkok. The other two provinces are rural, Buriram and Srisaket, located in the less developed northeastern region by the border of Cambodia. In each province, the survey is conducted in four randomly selected villages within a township. A detailed discussion of the survey design is provided in Samphantharak and Townsend (2010).

Our sample covers 710 households and 14 years of monthly data, starting in January 1999. The surveyed economies were exposed to a variety of aggregate and idiosyncratic shocks. For instance, rice cultivation is affected by seasonal variations in rainfall and temperature, restrictions on exports to the EU affected shrimp ponds, and milk cows' productivity fluctuates substantially across animals and over time. This rich set of shocks enables us to disentangle the roles of aggregate and idiosyncratic risk.

The TTMS collects detailed information on net business income, household assets and liabilities, and labor income. These data allow us to construct balance sheets at the household level and to compute the return on assets (ROA), defined as net business profits divided by business assets.[7] We also measure the fraction of entrepreneurs' wealth invested in the business relative to safe (real or financial) assets, which provides a measure of risk-taking. Finally, the data record household consumption, allowing us to characterize savings behavior and the role of non-business income. Appendix B details these variable constructions and discusses sources of measurement error and robustness checks.[8]

2.1 The determinants of expected entrepreneurial returns

Whether and how much to invest in a business depends on its risk–return trade-off. We study this trade-off by considering the required risk compensation implicit in entrepreneurial returns. Perfect risk sharing provides a natural benchmark: under this as-

[7] This measure excludes unrealized capital gains. As shown in Section 5, in our theory variation in expected returns is driven mainly by dividend yields rather than capital gains. Appendix B discusses in detail the implications of measurement error for our estimation.

[8] The household is the unit of observation. We treat the household as unitary; see Doepke and Tertilt (2016) for a discussion.

sumption, differences in expected returns are entirely driven by differences in exposure to aggregate risk, as idiosyncratic shocks can be perfectly insured or diversified. On the other extreme, entrepreneurs can be in financial autarky, without access to any insurance. In this case, only the total amount of risk is relevant to entrepreneurs, causing expected returns to vary with both aggregate and idiosyncratic volatility. When idiosyncratic shocks can be only partially insured or diversified, expected returns depend not only on the total variance of returns but also on the relative importance of each source of risk.

This discussion motivates a two-factor model to explain the cross-section of expected entrepreneurial returns based on exposure to aggregate and idiosyncratic shocks. This framework allows us to evaluate whether perfect risk sharing or autarky holds and, more importantly, whether a risk-based explanation can quantitatively account for the observed differences in entrepreneurial returns. We estimate this factor model following the two-pass regression methodology developed by Fama and MacBeth (1973).

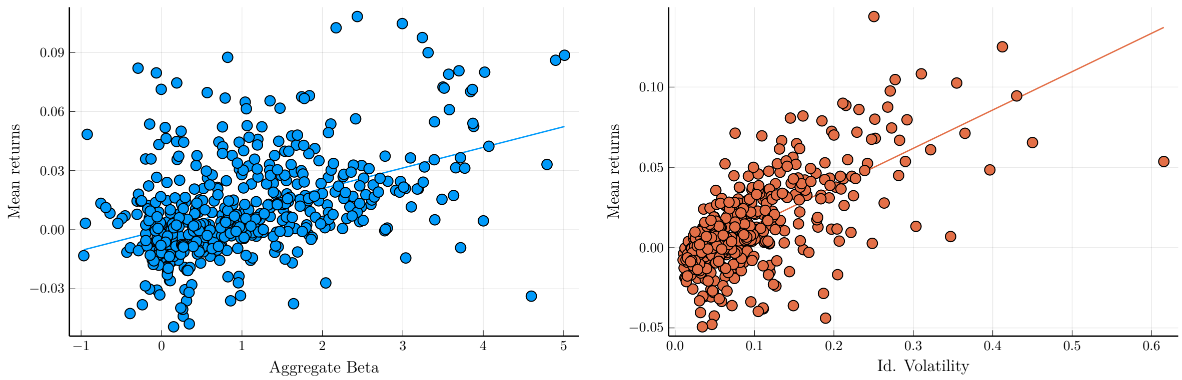
In the first stage, we estimate entrepreneurs' exposure to aggregate risk by running a time-series regression of returns for entrepreneur i, $R_{i,t}$, on the cross-sectional average return across all entrepreneurs in the same province, R_t^{agg}:

$$R_{i,t} = \alpha_i + \beta_i R_t^{agg} + \epsilon_{i,t}, \tag{1}$$

for each entrepreneur $i \in \{1, \ldots, N\}$. The variable R_t^{agg} plays the role of the market portfolio in standard CAPM tests. By averaging across entrepreneurs in a given region, idiosyncratic shocks are diversified, so β_i captures the exposure to aggregate risk for entrepreneur i.[9] We measure exposure to idiosyncratic risk by the variance of the residuals, $\sigma_i^2 = \text{var}[\epsilon_{i,t}]$. To address measurement concerns, in particular possible overfitting in the estimation of aggregate and idiosyncratic exposures, we report in Appendix B robustness checks using alternative measures based on shrinkage estimators.

[9] Following Samphantharak and Townsend (2018), we adopt the province as the relevant geographic unit. Moreover, we compute a leave-one-out mean of the returns, so that the return of entrepreneur i is regressed on the average return across all other entrepreneurs in the province.

Figure 1: Average returns vs. aggregate and idiosyncratic risk



Note: The left (right) panel shows a scatter plot of average time-series returns for each entrepreneur against aggregate beta (idiosyncratic volatility). Aggregate beta is measured as the slope of the time-series regression of individual returns on the leave-one-out average return in the entrepreneur's province. Idiosyncratic volatility is calculated as the volatility of residuals from the same regression. To limit the influence of outliers, we trim 1% of the observations in the left and right tails.

Figure 1 shows how the exposure to each source of risk is related to entrepreneurs' average return. We find a strong association between average returns and exposure to both aggregate and idiosyncratic risk. The left panel shows that entrepreneurs more exposed to aggregate risk have higher average returns. Similarly, the right panel shows a positive and significant relationship between average returns and idiosyncratic volatility.

Given that idiosyncratic volatility may be correlated with aggregate beta, the positive association between idiosyncratic risk and expected returns could still be consistent with perfect risk sharing. To address this point, we run a second-stage cross-sectional regression of the average return for entrepreneur i, $\overline{R}_i = \frac{1}{T}\sum_{t=1}^{T} R_{i,t}$, on the estimated exposures to aggregate risk, $\hat{\beta}_i$, and idiosyncratic risk, $\hat{\sigma}_i^2$:

$$\overline{R}_i = \lambda_0 + \lambda_{ag}\hat{\beta}_i + \lambda_{id}\hat{\sigma}_i^2 + u_i. \tag{2}$$

The coefficients λ_{ag} and λ_{id} correspond to the *prices of aggregate and idiosyncratic risk*, respectively. They capture the required compensation for one unit of exposure to each source of risk. The standard errors for the risk prices are obtained by embedding our two-stage procedure into a generalized method of moments (GMM) framework, which

Table 1: Cross-sectional regressions

Dependent Variable:	Mean ROA					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
(Intercept)	-0.007	0.009***	-0.003	-0.004	0.000	-0.003***
	(0.004)	(0.002)	(0.004)	(0.002)	(0.001)	(0.001)
Beta	0.02***		0.012***	0.009**	0.009***	0.008**
	(0.004)		(0.004)	(0.003)	(0.001)	(0.001)
Id. Variance		0.125***	0.103***	0.213***		
		(0.028)	(0.029)	(0.072)		
Id. Variance (PCA)					0.103	0.248**
					(0.066)	(0.126)
Fit statistics						
Observations	541	541	541	24	541	24
R^2	0.344	0.578	0.685	0.892	0.703	0.876
Adjusted R^2	0.343	0.577	0.684	0.882	0.701	0.858

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The table shows a cross-sectional regression of average returns on the aggregate beta or idiosyncratic variance (or both). Standard errors account for the uncertainty in estimating the regressors and are robust to contemporaneous correlations across entrepreneurs. Columns (5) and (6) include the exposure to the latent factor as an additional control.

accounts for the sampling uncertainty in $\hat{\beta}_i$ and $\hat{\sigma}_i^2$.[10]

Table 1 reports the second-stage results. We find a positive and significant price of risk for both β_i and σ_i^2, and these two factors together explain a substantial fraction of the cross-sectional variation in expected returns. Column 1 shows that the CAPM-inspired one-factor model already explains 34% of the variation in average returns, a surprisingly strong result given the limited empirical success of the CAPM in equity markets (see, e.g., Fama and French 1992). Column 2 shows that idiosyncratic risk explains an even larger fraction, with an adjusted R^2 of 58%. Together, the two factors account for 68% of the variation, close to the 73% R^2 of the three-factor Fama–French model applied to 25 equity portfolios sorted by size and book-to-market. This high explanatory power is particularly striking given that our evidence comes from more than 500 entrepreneurs in a developing-country setting.

An important concern is that our risk exposure measures are potentially noisy estimates of the true betas and idiosyncratic volatilities, which may bias our results. To miti-

[10]See Cochrane (2009) for a discussion on how to correct the standard errors in two-pass regressions.

gate this issue, we follow the standard empirical asset pricing approach (see, e.g., Black et al. 1972) and group entrepreneurs into portfolios using a 5×5 double sort on aggregate beta and idiosyncratic risk.[11] This procedure reduces measurement error in exposures and yields more reliable cross-sectional estimates.

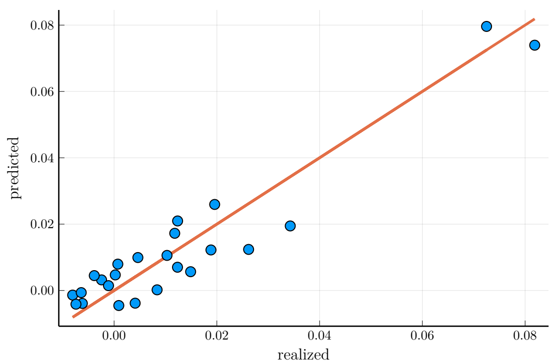
Column 4 reports the portfolio-level results. We obtain a similar price of aggregate risk and a larger price of idiosyncratic risk, consistent with attenuation bias at the individual level.[12] The two risk factors together explain 88% of the variation in group-level returns. Figure 2 illustrates the fit of the two-factor model by comparing predicted and realized average returns. The model successfully accounts for the substantial cross-sectional variation in expected returns observed in our data. Formally, we cannot reject that the data are generated by the two-factor model, and the remaining residual variation is consistent with sampling noise. Both risk factors are necessary to obtain this fit: in particular, we can reject the hypothesis that a single-factor model based on either aggregate or idiosyncratic risk alone suffices. Taken together, the evidence points to a two-factor structure as necessary to explain the cross-section of entrepreneurial returns.

Dealing with omitted factors. The success of the two-factor model relies on the assumption that we capture the relevant aggregate sources of risk. If entrepreneurs are in fact exposed to multiple latent aggregate factors, then the error term in Eq. (1) may reflect not only idiosyncratic shocks but also omitted aggregate components. In this case, both estimated risk exposures and the prices of risk could be biased.[13] Building on recent advances in the estimation of asset-pricing models, we extend our framework to allow explicitly for latent aggregate factors.

[11] Although these portfolios are not tradeable, forming groups reduces estimation error and allows us to capture more precisely the joint relationship between risk exposures and average returns. To avoid within-portfolio diversification, we assign to each portfolio the average of entrepreneurs' idiosyncratic volatilities rather than the volatility of the portfolio-level average return.

[12] Appendix B.3 shows that applying the shrinkage estimator of Ledoit and Wolf (2004) to the idiosyncratic variance across the full cross-section of entrepreneurs yields similar results.

[13] See, e.g., Burmeister and McElroy (1988) for a discussion of biases caused by omitted factors.



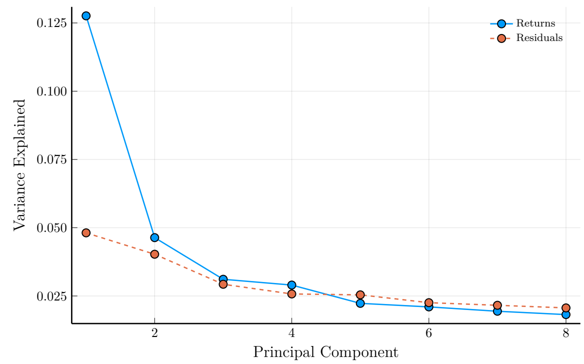


Figure 2: Realized vs. predicted returns

Figure 3: Scree plot

Note: The left panel shows the 45° line and a scatter plot of the predicted returns of the two-factor model from column (4) of Table 1 and actual average returns at the portfolio level. The right panel shows the scree plot of the eigenvalues of the covariance matrix of returns and of the residuals from the first-stage regression.

Consider the generalized model in which entrepreneurial returns are exposed to a vector of aggregate factors $f_t \in \mathbb{R}^K$:

$$R_{i,t} = \mu_i + \beta_i^\top f_t + \epsilon_{i,t}, \qquad \epsilon_{i,t} \sim \mathcal{N}(0, \sigma_i^2), \quad f_t \perp \epsilon_{i,t},$$

where f_t has mean zero and μ_i denotes expected excess returns. The relationship between expected returns and risk exposures is then:

$$\mu_i = \lambda_0 + \lambda_{ag}^\top \beta_i + \lambda_{id} \sigma_i^2 + u_i. \tag{3}$$

Here, $\lambda_{ag} \in \mathbb{R}^K$ is the vector of risk prices for the aggregate factors, so $\lambda_{ag}^\top \beta_i$ represents compensation for aggregate exposures.[14] The coefficient $\lambda_{id} \in \mathbb{R}$ captures the price of idiosyncratic risk. Finally, the residual term $u_i \sim \mathcal{N}(0, \sigma_u^2)$ absorbs any remaining misspecification, such as measurement error or non-risk-based determinants of returns.

Importantly, some aggregate factors may be latent, i.e., $f_t = [f_{o,t}^\top, f_{u,t}^\top]^\top$, where $f_{o,t}$ is observed and $f_{u,t}$ is unobserved. In this case, realized returns can be written as:

$$R_{i,t} = \mu_i + \beta_{i,o}^\top f_{o,t} + \tilde{\epsilon}_{i,t}, \tag{4}$$

[14] Appendix B.5 shows that our formulation also accommodates time-varying loadings and risk premia.

where $\tilde{\epsilon}_{i,t} = \beta_{i,u}^{\top} f_{u,t} + \epsilon_{i,t}$. If $f_{u,t}$ is correlated with $f_{o,t}$, omitted factors bias the estimated exposure to observable risk, $\beta_{i,o}$. Moreover, differences in the residual variance, $\mathrm{Var}[\tilde{\epsilon}_{i,t}]$, would reflect not only idiosyncratic shocks but also exposure to latent factors. As a result, both estimated betas and the price of idiosyncratic risk in the second stage may be biased.

When omitted factors represent an important fraction of the variation in entrepreneurial returns, the residuals of regression (4) will have a factor structure. The omitted factors would drive correlation across residuals. Up to a rotation, the loadings on latent factors can be recovered using the principal components of the residuals. Given these loadings, we can control for the exposure to latent factors in the second-stage regression.

Building on this observation, Giglio and Xiu (2021) propose a method to estimate factor models with latent components in settings with large N and T. Hence, their methodology is particularly well-suited to settings with a sizable cross-section such as ours.[15] Their three-pass procedure consists of: (i) time-series regressions of returns on observable factors, (ii) principal components analysis (PCA) of the residuals to extract latent factors, and (iii) cross-sectional regressions of average returns on the estimated exposures to both observable and latent factors. This approach delivers consistent estimates of risk prices and is robust to factor rotation. We implement this methodology to assess the empirical relevance of latent aggregate factors, with estimation details provided in Appendix B.5.

We start by selecting the number of latent factors. Figure 3 shows the scree plot, which reports the share of variance explained by each principal component, for both returns and residuals of the first-stage regression.[16] The observable factor corresponds to the average return in the region. While returns display the familiar "elbow" pattern in the scree plot, the residuals do not, suggesting that the observable factor accounts for nearly all common variation. Formally, we implement the test of Onatski (2009) for the number of latent

[15]The main empirical application in Giglio et al. (2021) uses $N = 647$, comparable to our $N = 541$. Giglio and Xiu (2021) also show in Monte Carlo simulations that the estimator performs well with modest sample sizes ($N = 200$, $T = 120$).

[16]Given our unbalanced panel, we apply an iterative procedure to compute principal components (see, e.g., Stock and Watson 2002). Appendix B.5 provides details.

factors and cannot reject the null hypothesis of zero omitted factors.

Nevertheless, we allow for a single latent factor to assess sensitivity to potential misspecification. Column 5 of Table 1 reports the three-pass regression results. Following Giglio et al. (2021), we compute standard errors using a wild bootstrap, as recommended for the three-pass procedure. In this specification, average returns are regressed on exposures to both observable and latent factors, while our measure of idiosyncratic risk is purged of latent-factor influence. The results are nearly identical to those in column 3. Column 6 shows the corresponding portfolio-level analysis, which closely matches the results in column 4. In both cases, we cannot reject that expected returns are fully explained by exposure to aggregate and idiosyncratic risk.

Overall, these findings suggest that our residual variance measure primarily captures idiosyncratic shocks rather than omitted factors. The strong fit of the two-factor model further indicates that non-risk-based explanations, represented by u_i in Eq. (3), play at most a minor role in our setting.

Appendix B documents additional robustness checks. We show that our factor model accommodates time variation in factor loadings and prices of risk, discuss the role of classical measurement error in returns, and demonstrate that our main results are preserved when we use shrinkage estimators for idiosyncratic variance, betas, and expected returns.

Risk and return decomposition. To better understand the role of each risk factor, we decompose entrepreneurial returns into aggregate and idiosyncratic components. Table 2 reports the decomposition of the risk premium and variance of returns based on the results from column 4 of Table 1. Most of the risk is idiosyncratic, accounting for 94% of the variance, while idiosyncratic shocks explain nearly half of the risk premium. The Sharpe ratio of the aggregate component is almost five times larger than that of the idiosyncratic component.[17]

[17] Samphantharak and Townsend (2018) compute a similar decomposition using total portfolio returns, including risky and safe assets. Our focus is on the risky component, which is the relevant one for mapping

Table 2: Aggregate and idiosyncratic components of risk and return

	Risk premium	% of returns	Volatility	% of variance	Sharpe ratio
Total returns	4.4%	100%	21.0%	100%	0.21
Aggregate component	2.4%	54.7%	5.1%	6.0%	0.47
Idiosyncratic component	2.0%	45.3%	20.3%	94.0%	0.10

Note: The aggregate component is measured as the coefficient on aggregate beta in the cross-sectional regression times the average beta across entrepreneurs. The idiosyncratic component is measured as the coefficient on idiosyncratic risk times the idiosyncratic variance averaged across entrepreneurs. The variance decomposition is computed at the individual level based on the first-stage regression and then averaged across all entrepreneurs.

The results in Table 2 are inconsistent with both perfect risk sharing and financial autarky. The fact that the idiosyncratic risk premium accounts for almost half of total expected returns indicates substantial limits to diversification, allowing us to reject perfect risk sharing. Financial autarky, by contrast, implies that the Sharpe ratio of the idiosyncratic component should exceed that of the aggregate component in proportion to their volatilities, as shown in Appendix C.2.[18] In our data, however, the opposite holds: the aggregate Sharpe ratio is substantially higher. This pattern is consistent with partial insurance, which reduces the effective exposure to idiosyncratic shocks. We revisit this decomposition through the lens of the structural model in Section 6.

2.2 Risk-taking and savings over the life cycle

We now turn to entrepreneurs' risk-taking and savings decisions. The choice of how much to invest in the business depends on the risk and return of the activity as well as households' risk appetite. Behavior towards risk may vary systematically with household characteristics such as age and family size. We therefore study how risk-taking and consumption decisions depend on expected returns and a range of demographic controls.

To ensure stationarity, we scale all variables by financial wealth (i.e., net worth) and focus on ratios rather than levels. Risk-taking is measured by the ratio of business value

into the structural model in Section 3.

[18] In a two-factor model, expected returns are $p^{ag}\sigma_{ag} + p^{id}\sigma_{id}$, where (p^{ag}, p^{id}) are prices of risk and $(\sigma_{ag}, \sigma_{id})$ are volatilities. Perfect risk sharing corresponds to $p^{id} = 0$. Under autarky, the Sharpe ratios satisfy $p^{ag} = \gamma\sigma_{ag}$ and $p^{id} = \gamma\sigma_{id}$, implying that the ratio of Sharpe ratios equals the ratio of volatilities.

Table 3: Risk-taking and consumption behavior over the life cycle

	Risk-taking			Consumption-wealth ratio		
	(1)	(2)	(3)	(4)	(5)	(6)
Age group: 1	0.30***	0.31***	0.27***	0.14***	0.14***	0.10***
	(0.01)	(0.01)	(0.02)	(0.009)	(0.008)	(0.007)
Age group: 2	0.27***	0.28***	0.25***	0.13***	0.13***	0.11***
	(0.01)	(0.01)	(0.01)	(0.009)	(0.007)	(0.007)
Age group: 3	0.27***	0.27***	0.25***	0.12***	0.12***	0.11***
	(0.01)	(0.01)	(0.01)	(0.008)	(0.007)	(0.007)
Age group: 4	0.23***	0.23***	0.22***	0.10***	0.10***	0.08***
	(0.02)	(0.01)	(0.01)	(0.007)	(0.006)	(0.006)
Age group: 5	0.22***	0.22***	0.21***	0.11***	0.11***	0.09***
	(0.01)	(0.01)	(0.01)	(0.008)	(0.008)	(0.007)
Year FE		Yes	Yes		Yes	Yes
Additional controls			Yes			Yes
N	8,680	8,680	6,499	8,548	8,548	6,495
Adjusted R^2	0.015	0.025	0.139	0.010	0.020	0.070

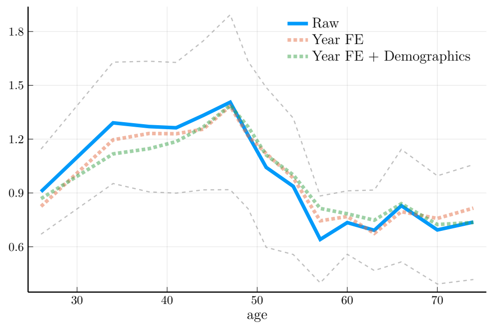
Clustered (year & household) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: Age effects correspond to the conditional expectation of the outcome variable evaluated at the mean of the continuous controls and averaged over the year, province, and sector fixed effects (when included). Additional controls include dummies for province, household size, number of children, sector affiliation, the entrepreneur's average return on the business, and exposures to aggregate and idiosyncratic risk.

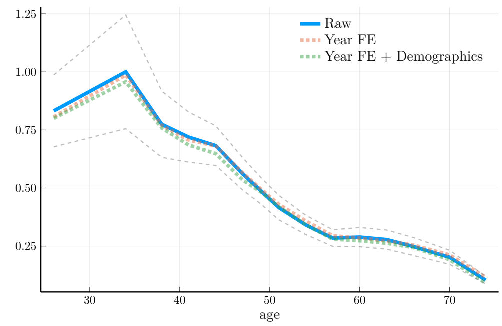
to financial wealth, and savings behavior by the ratio of consumption to financial wealth. Using ratios mitigates the influence of aggregate shocks and captures more stable life-cycle relationships. Entrepreneurs are grouped into five age categories, from 25 to 80 years, with cutoffs chosen to ensure roughly equal sample sizes. Appendix B.6 shows that the main patterns are robust to alternative specifications with more age groups. Standard errors are clustered by household and year to allow for correlated shocks within households over time and across households within a given year. We aggregate the data to the annual level to isolate age effects and trim 1% of observations in both tails to reduce the impact of outliers.

We compute age effects controlling for year fixed effects and additional covariates following the methodology of Kaplan (2012). Table 3 presents the results. We find that risk-taking displays a strong life-cycle pattern. Column 1, without additional controls, shows that the share of wealth invested in the business declines sharply with age, with the coef-

Figure 4: Labor Income and Human-Financial Ratio



(a) Labor Income



(b) Human-financial wealth ratio

Note: Labor income is normalized by the average labor income across all entrepreneurs. Human wealth is the present discounted value of future labor income, as measured by average income for each age. Grey dashed lines represent the 95% confidence interval for the raw estimates, i.e., the ones without any controls.

ficient for the oldest group roughly 30% smaller than that for the youngest. Adding time fixed effects in column 2 does not change the results, suggesting that aggregate shocks affect both business and financial wealth but leave their ratio roughly unchanged. Column 3 includes demographic and geographic controls, as well as risk/return controls such as household size, number of children, province, sector affiliation, average business return, and risk exposures. Even with these controls, we find substantial variation in risk-taking across age groups. Since these patterns are not explained by differences in risk or expected returns, they are consistent with systematic life-cycle variation in risk attitudes.

The consumption-wealth ratio also exhibits a strong life-cycle profile. Unlike risk-taking, which declines monotonically with age, the consumption-wealth ratio follows a U-shape: it decreases at first and then rises again for the oldest group. Appendix B.6 reports formal tests that statistically confirm these life-cycle patterns.

Both outcomes are potentially influenced by outside income. The left panel of Figure 4 shows the life-cycle profile of labor income received by entrepreneurial households. The right panel plots the ratio of human wealth—the present discounted value of expected future labor income—to financial wealth. This human-financial ratio declines with age, and will play an important role in rationalizing the patterns in Table 3.

We draw three conclusions from these results. First, risk-taking and consumption both exhibit substantial life-cycle variation. Second, the decline in risk-taking with age is not explained by differences in risk or returns, but is consistent with differences in effective risk tolerance across households. Third, the stability of results across specifications suggests that year fixed effects and aggregate shocks do not materially affect the life-cycle patterns. Taken together, these findings—along with the evidence on partial insurance from Section 2.1—form the basis for the theoretical model we develop below.

3 A life-cycle model of entrepreneurial risk taking

We develop a life-cycle model of entrepreneurial activity with two main ingredients: (i) imperfect insurance of idiosyncratic risk and (ii) finite lives with imperfect altruism. These features are crucial to account for the empirical patterns documented in Section 2.

3.1 Environment

Time is continuous, and the economy is populated by entrepreneurs and wage earners. Population grows at rate g, and the share of entrepreneurs in the population is constant and given by χ_e. The sets of entrepreneurs and wage earners alive at period t are denoted by $\mathcal{E}_t$ and $\mathcal{W}_t$, respectively. Entrepreneurs live for T periods and leave bequests to their offspring, while wage earners are modeled as infinitely lived for simplicity. Appendix D.2 shows our results are unchanged when workers also have finite lives. All households receive labor income, but only entrepreneurs have access to a production technology. Production is subject to both aggregate and idiosyncratic shocks. Markets for aggregate risk are complete, so households can buy and sell insurance against aggregate shocks in a frictionless market. By contrast, insurance against idiosyncratic shocks is imperfect. In addition, households can borrow and lend at the riskless rate r_t.

Technology. Entrepreneur i combines capital $k_{i,t}$ and hired labor $l_{i,t}$ to produce the final good $\tilde{y}_{i,t}$, which is the numeraire in this economy:

$$\tilde{y}_{i,t} = A_t k_{i,t}^{\alpha} l_{i,t}^{1-\alpha}. \tag{5}$$

We denote scaled output by $y_{i,t} = \tilde{y}_{i,t}/A_t$. Throughout the paper, variables that grow with aggregate productivity A_t are denoted with a tilde, while their stationary counterparts are written without it.

Aggregate productivity follows a geometric Brownian motion,

$$\frac{dA_t}{A_t} = \mu_A\, dt + \sigma_A\, dZ_t, \tag{6}$$

where Z_t is a standard Brownian motion.

Entrepreneurs can adjust their capital stock either by investing in new capital or by purchasing existing capital from other entrepreneurs. Investment is risky and subject to adjustment costs. Given a total investment of $\iota_{i,t} A_t k_{i,t}$, the law of motion for capital is

$$\frac{dk_{i,t}}{k_{i,t}} = (\Phi(\iota_{i,t}) - \delta)\, dt + \sigma_I\, dZ_{i,t}, \tag{7}$$

where $Z_{i,t}$ is an idiosyncratic Brownian motion, independent across entrepreneurs. The investment function $\Phi(\cdot)$ satisfies $\Phi(0) = 0$, $\Phi'(\cdot) > 0$, and $\Phi''(\cdot) < 0$, where the concavity captures the presence of adjustment costs. Entrepreneurs may also adjust their capital stock by trading at the market price $\tilde{q}_t = q_t A_t$, so the market value of the business is $\tilde{q}_t k_{i,t}$. In equilibrium, q_t is non-stochastic, implying that the relative price of capital $\tilde{q}_t$ moves proportionally with aggregate productivity.

Aggregate capital, $k_t = \int_{\mathcal{E}_t} k_{i,t}\, di$, is unaffected by idiosyncratic shocks, which diversify in the aggregate:

$$dk_t = \left[\int_{\mathcal{E}_t} (\Phi(\iota_{i,t}) - \delta)\, k_{i,t}\, di \right] dt. \tag{8}$$

The return on entrepreneurial projects is the sum of dividend yield and capital gains:

$$\begin{aligned} dR_{i,t} &= \frac{\tilde{y}_{i,t} - \tilde{w}_t l_{i,t} - \iota_{i,t} A_t k_{i,t}}{\tilde{q}_t k_{i,t}} dt + \frac{d(\tilde{q}_t k_{i,t})}{\tilde{q}_t k_{i,t}} \\ &\equiv \mu^R_{i,t} dt + \sigma_A dZ_t + \sigma_I dZ_{i,t}, \end{aligned}$$

where $\tilde{w}_t = w_t A_t$ denotes the wage and $\mu^R_{i,t}$ the expected return.

Applying Itô's lemma to compute expected capital gains, the expected return is

$$\mu^R_{i,t} \equiv \frac{y_{i,t} - w_t l_{i,t} - \iota_{i,t} k_{i,t}}{q_t k_{i,t}} + \frac{\dot{q}_t}{q_t} + \mu_A + \Phi(\iota_{i,t}) - \delta. \tag{9}$$

Preferences. Entrepreneurs live for T periods and have isoelastic preferences over consumption $\tilde{c}_{i,t}$ with curvature parameter γ. They also derive utility from leaving bequests:

$$\mathbb{E}_{s_i}\left[\int_{s_i}^{s_i+T} e^{-\rho(t-s_i)} \frac{\tilde{c}_{i,t}^{1-\gamma}}{1-\gamma} dt + e^{-\rho(T-s_i)} (1-\psi)^{\gamma} V^* \frac{\tilde{n}_{i,s_i+T}^{1-\gamma}}{1-\gamma}\right], \tag{10}$$

where $\tilde{n}_{i,t}$ denotes financial wealth (net worth) and s_i is the entrepreneur's birthdate.

The parameter ψ captures the strength of the bequest motive. When $\psi = 1$, entrepreneurs give no weight to offspring, while $\psi = 0$ corresponds to the infinite-horizon case.[19] Intermediate values $0 < \psi < 1$ represent imperfect altruism.

In addition to business income, entrepreneurial households receive labor income, consistent with the data showing that households have multiple income sources. Labor supply is denoted by $\bar{l}_{i,t}$ and varies deterministically over the life cycle.[20]

Financial frictions. Entrepreneurs face a dynamic moral hazard problem, similar to He and Krishnamurthy (2012) and Di Tella (2017). Following Di Tella (2017), we restrict attention to short-term contracts. Aggregate shocks are publicly observable, while idiosyn-

[19] V^*, defined in Appendix A.1, is the value-function coefficient for an infinite-horizon agent.
[20] $\bar{l}_{i,t}$ denotes exogenous household labor supply, while $l_{i,t}$ is labor hired in production.

cratic shocks are observed only by the entrepreneur. An entrepreneur may divert capital, but a fraction $1-\phi$ of diverted resources is lost, where $\phi \in (0,1)$ captures the severity of the moral hazard problem.

As in the dynamic contracting literature, the optimal allocation can be implemented through a market structure with (i) aggregate insurance, (ii) idiosyncratic insurance, and (iii) a riskless asset yielding r_t (see Appendix E for a formal derivation). Because aggregate shocks are verifiable, aggregate risk sharing is unrestricted, while idiosyncratic insurance is limited by incentive constraints. Entrepreneur i pays $p_t^{ag}\tilde{\theta}_{i,t}^{ag}$ to reduce aggregate volatility by $\tilde{\theta}_{i,t}^{ag}$, where p_t^{ag} is the price of aggregate insurance. Idiosyncratic insurance $\tilde{\theta}_{i,t}^{id}$ is costless in equilibrium – since providers diversify across entrepreneurs – but bounded by the *skin-in-the-game* constraint:

$$\tilde{\theta}_{i,t}^{id} \leq (1-\phi)\,\tilde{q}_t k_{i,t}\sigma_I. \tag{11}$$

This market structure represents one possible implementation of the optimal contract. Rather than formal contracts, it may reflect informal insurance through kinship networks, as documented by Kinnan and Townsend (2012). Alternatively, it can be viewed as an equity constraint that prevents entrepreneurs from fully selling claims to outside investors, as in Chen et al. (2010) and Panousi and Papanikolaou (2012). Section 6 presents an alternative implementation based on flexible lending contracts.

Constraint (11) binds in equilibrium, so entrepreneurs are *insurance-constrained*. To isolate entrepreneurial risk, we initially abstract from borrowing frictions, which we revisit in Section 5.2. Households face only the natural borrowing limit:

$$\tilde{n}_{i,t} \geq -\tilde{h}_{i,t}. \tag{12}$$

Entrepreneurs' problem. An entrepreneur i of age a_i chooses, taking prices $(\tilde{q}, \tilde{w}, r, p^{ag})$ as given, stochastic processes $(\tilde{c}_i, \tilde{\theta}_i^{ag}, \tilde{\theta}_i^{id}, k_i, l_i, \iota_i)$ to solve

$$\tilde{V}_t(\tilde{n}_i, a_i) = \max_{\tilde{c}_i, \tilde{\theta}_i^{ag}, \tilde{\theta}_i^{id}, k_i, l_i, \iota_i} \mathbb{E}_t \left[\int_0^{T-a_i} e^{-\rho z} \frac{\tilde{c}_{i,t+z}^{1-\gamma}}{1-\gamma} dz + e^{-\rho(T-a_i)} (1-\psi)^{\gamma} V^* \frac{\tilde{n}_{i,t+T-a_i}^{1-\gamma}}{1-\gamma} \right], \quad (13)$$

subject to (11), (12), $c_{i,t}, k_{i,t} \geq 0$, and the law of motion for financial wealth:

$$\begin{aligned} d\tilde{n}_{i,t} = & \left[(\tilde{n}_{i,t} - \tilde{q}_t k_{i,t}) r_t + \tilde{q}_t k_{i,t} \mu_{i,t}^R - p_t^{ag} \tilde{\theta}_{i,t}^{ag} + \tilde{w}_t \bar{l}_{i,t} - \tilde{c}_{i,t} \right] dt \\ & + \left(\tilde{q}_t k_{i,t} \sigma_A - \tilde{\theta}_{i,t}^{ag} \right) dZ_t + \left(\tilde{q}_t k_{i,t} \sigma_I - \tilde{\theta}_{i,t}^{id} \right) dZ_{i,t}, \end{aligned}$$

given initial wealth $\tilde{n}_{i,t} = \tilde{n}_i > -\tilde{h}_{i,t}$.

The first bracketed term captures the expected change in wealth. Entrepreneurs allocate $\tilde{n}_{i,t} - \tilde{q}_t k_{i,t}$ to the riskless asset (yield r_t) and $\tilde{q}_t k_{i,t}$ to the risky technology (expected return $\mu_{i,t}^R$). They pay $p_t^{ag} \tilde{\theta}_{i,t}^{ag}$ for aggregate insurance, receive labor income $\tilde{w}_t \bar{l}_{i,t}$, and consume $\tilde{c}_{i,t}$. The last two terms represent net exposures to aggregate and idiosyncratic shocks after insurance.

Wage earners' problem. Wage earners do not have access to a production technology and, for simplicity, are assumed to live infinitely. Appendix D.2 shows that allowing for finite lives with bequests, or an endogenous occupational choice, does not alter our main results. Wage earners and entrepreneurs share isoelastic period utility with curvature parameter γ,[21] but differ in their discount rates, as is standard in models with heterogeneous returns (e.g., Kiyotaki and Moore 1997).

[21] As shown by Swanson (2012), and consistent with Lemma 1 below, the coefficient of relative risk aversion in the presence of labor is affected by, but not equal to, γ.

The problem of a wage earner $j \in \mathcal{W}_t$ is

$$\tilde{V}_t^w(\tilde{n}_j) = \max_{\tilde{c}_j, \tilde{\theta}_j^{ag}} \mathbb{E}_t \left[\int_t^\infty e^{-\rho_w(z-t)} \frac{\tilde{c}_{j,z}^{1-\gamma}}{1-\gamma} dz \right], \tag{14}$$

subject to $\tilde{c}_{j,t} \geq 0$, $\tilde{n}_{j,t} \geq -\tilde{h}_{j,t}$, where $\tilde{h}_{j,t}$ denotes human wealth, and financial wealth evolves as

$$d\tilde{n}_{j,t} = \left[\tilde{n}_{j,t} r_t - p_t^{ag} \tilde{\theta}_{j,t}^{ag} + \tilde{w}_t \bar{l}_{j,t} - \tilde{c}_{j,t}\right] dt - \tilde{\theta}_{j,t}^{ag} dZ_t,$$

with initial wealth $\tilde{n}_{j,t} = \tilde{n}_j > -\tilde{h}_{j,t}$.

The insurance position $\tilde{\theta}_{j,t}^{ag}$ can be positive or negative, so wage earners may either purchase aggregate insurance or provide it to entrepreneurs.[22]

Equilibrium.

Definition 1. *A competitive equilibrium consists of aggregate stochastic processes for the capital stock k, the interest rate r, the wage $\tilde{w}$, the relative price of capital $\tilde{q}$, and the price of aggregate insurance p^{ag}; individual processes for entrepreneurs $i \in \mathcal{E}_t$ — consumption $\tilde{c}_i$, financial wealth $\tilde{n}_i$, capital k_i, labor l_i, investment ι_i, aggregate insurance $\tilde{\theta}_i^{ag}$, and idiosyncratic insurance $\tilde{\theta}_i^{id}$; and for wage earners $j \in \mathcal{W}_t$ — consumption $\tilde{c}_j$ and aggregate insurance $\tilde{\theta}_j^{ag}$, such that:*

(a) *The aggregate capital stock evolves according to (8), given initial k_0.*

(b) *Entrepreneurs' allocations $(\tilde{c}_i, \tilde{\theta}_i^{ag}, \tilde{\theta}_i^{id}, k_i, l_i, \iota_i)$ solve (13), given prices $(\tilde{q}, \tilde{w}, r, p^{ag})$.*

(c) *Wage earners' allocations $(\tilde{c}_j, \tilde{\theta}_j^{ag})$ solve (14), given prices $(\tilde{w}, r, p^{ag})$.*

(d) *Markets clear for all $t \geq 0$:*

[22] Wage earners also provide idiosyncratic insurance. Since they can diversify idiosyncratic shocks across households, their wealth is only exposed to aggregate risk.

i. ***Goods:***

$$\int_{\mathcal{E}_t} \tilde{c}_{i,t} \, di + \int_{\mathcal{W}_t} \tilde{c}_{j,t} \, dj + \int_{\mathcal{E}_t} \iota_{i,t} A_t k_{i,t} \, di = \int_{\mathcal{E}_t} \tilde{y}_{i,t} \, di.$$

ii. ***Capital and labor:***

$$\int_{\mathcal{E}_t} k_{i,t} \, di = k_t, \qquad \int_{\mathcal{E}_t} l_{i,t} \, di = \int_{\mathcal{E}_t} \bar{l}_{i,t} \, di + \int_{\mathcal{W}_t} \bar{l}_{j,t} \, dj.$$

iii. ***Aggregate insurance:***

$$\int_{\mathcal{E}_t} \tilde{\theta}^{ag}_{i,t} \, di + \int_{\mathcal{W}_t} \tilde{\theta}^{ag}_{j,t} \, dj = 0.$$

iv. ***Riskless bond:***

$$\int_{\mathcal{E}_t} \left(\tilde{n}_{i,t} - \tilde{q}_t k_{i,t} \right) di + \int_{\mathcal{W}_t} \tilde{n}_{j,t} \, dj = 0.$$

3.2 Solution to entrepreneurs' problem

We now describe the solution to the entrepreneurs' problem in a stationary equilibrium, where scaled aggregate variables are constant, i.e. $w_t = w$ and $q_t = q$.

Maximizing expected returns. Entrepreneurs choose labor demand $l_{i,t}$ and the investment rate $\iota_{i,t}$ to maximize the expected return on the business, given in (9). The first-order condition for labor demand equates the wage to the marginal product of labor:

$$w = (1 - \alpha) \left(\frac{k_{i,t}}{l_{i,t}} \right)^{\alpha}. \tag{15}$$

Hence the capital–labor ratio is equalized across entrepreneurs and coincides with the aggregate capital–labor ratio $K_t \equiv k_t / \bar{l}_t$, where k_t is the aggregate capital stock and $\bar{l}_t$ the aggregate labor supply. In a stationary equilibrium, capital grows at the same rate g as the population.

Similarly, the first-order condition for investment implies

$$\Phi'(\iota_{i,t}) = \frac{1}{q} \quad \Rightarrow \quad \iota_{i,t} = (\Phi')^{-1}\left(\tfrac{1}{q}\right) \equiv \iota(q), \tag{16}$$

where $\iota(q)$ is increasing in q by concavity of $\Phi(\cdot)$. Intuitively, entrepreneurs trade off higher expected capital gains (through $\Phi'(\iota)$) against lower dividend yields ($1/q$).

Substituting (15) and (16) into (9), expected returns simplify to

$$\mu^R = \frac{\alpha K^{\alpha-1} - \iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta, \tag{17}$$

which is identical across entrepreneurs in a stationary equilibrium: $\mu^R_{i,t} = \mu^R$. Heterogeneity in realized returns arises from exposure to aggregate and idiosyncratic shocks.

Human and total wealth. The relevant notion of wealth in this environment is *total wealth*, $\omega_{i,t} \equiv n_{i,t} + h_{i,t}$, the sum of financial and human wealth. The lemma below shows that the entrepreneur's value function depends only on total wealth and age. All proofs are provided in Appendix A.

Lemma 1. *Suppose the economy is in a stationary equilibrium. Human wealth satisfies*

$$\frac{\partial h(a)}{\partial a} = (r + p^{ag}\sigma_A - \mu_A)\, h(a) - w\bar{l}(a), \quad h(T) = 0. \tag{18}$$

The (scaled) value function is[23]

$$V(n,a) = \zeta(a)^{-\frac{1}{\gamma}} \frac{(n + h(a))^{1-\gamma}}{1-\gamma}, \tag{19}$$

where $\zeta(a)$ is the consumption-total wealth ratio. The effective risk aversion is

$$-\frac{V_{nn} n}{V_n} = \frac{\gamma}{1 + h(a)/n}. \tag{20}$$

[23]The scaled value function relates to the original value function by $\tilde{V}_t(\tilde{n}, a) = A_t^{1-\gamma} V(\tilde{n}/A_t, a)$.

Equation (18) shows that human wealth is the present discounted value of future labor income, where the discount rate includes the aggregate risk premium $p^{ag}\sigma_A$. Because wages co-move with aggregate productivity, human wealth is effectively a risky asset. This mechanism is consistent with Benzoni et al. (2007), who show that labor income becomes highly correlated with stock returns when it is cointegrated with output. Since human wealth depends only on age, we drop the household index and write $h_{i,t} = h(a_{i,t})$.

The second part of the lemma shows that the value function is CRRA in total wealth, with an age-dependent consumption-wealth ratio. Importantly, effective risk aversion is decreasing in the human-financial wealth ratio $h_{i,t}/n_{i,t}$. This provides a natural channel for life-cycle variation in risk-taking, as this ratio declines with age in the data.

Policy functions. The next proposition characterizes entrepreneurs' optimal choices in the stationary equilibrium.

Proposition 1. *Suppose the economy is in a stationary equilibrium. Then:*

i) *Capital demand satisfies*

$$\frac{qk_{i,t}}{n_{i,t}} = \frac{1 + \frac{h_{i,t}}{n_{i,t}}}{\gamma} \frac{p^{id}}{\phi\sigma_I}, \tag{21}$$

where p^{id}, the shadow price of idiosyncratic insurance, is

$$p^{id} = \frac{\mu^R - r - p^{ag}\sigma_A}{\phi\sigma_I}. \tag{22}$$

ii) *The price of aggregate insurance is $p^{ag} = \gamma\sigma_A$, and entrepreneurs' aggregate insurance position is*

$$\theta^{ag}_{i,t} = (qk_{i,t} - n_{i,t})\sigma_A. \tag{23}$$

iii) *The consumption-wealth ratio is*

$$\frac{c_{i,t}}{n_{i,t}} = \frac{\bar{r}}{1 - \psi e^{-\bar{r}(T-a_i)}} \left(1 + \frac{h_{i,t}}{n_{i,t}}\right), \tag{24}$$

where $\bar{r} \equiv \frac{1}{\gamma}\rho + \left(1 - \frac{1}{\gamma}\right)\left[r + \frac{(p^{ag})^2 + (p^{id})^2}{2\gamma}\right]$.

The demand for capital depends on three terms: effective risk aversion, $\gamma/(1 + \frac{h_{i,t}}{n_{i,t}})$, the shadow price of idiosyncratic insurance, p^{id}, and the undiversifiable component of risk, $\phi\sigma_I$. Because effective risk aversion is shaped by the human-financial wealth ratio, cross-sectional differences in risk-taking inherit the life-cycle patterns of this ratio.

Equation (21) shows that capital demand depends only on the price and quantity of *idiosyncratic* risk. Because aggregate insurance is freely traded, entrepreneurs can separate the business scale decision from their aggregate risk exposure. Investment in the business thus boils down to a choice about how much idiosyncratic risk to bear.

Equation (22) highlights that p^{id} is the Sharpe ratio of an entrepreneur who is fully insured against aggregate risk. Since this Sharpe ratio is positive in equilibrium, the skin-in-the-game constraint is always binding, implying $\theta^{id}_{i,t} = (1 - \phi)q_t k_{i,t}\sigma_I$. Entrepreneurs therefore purchase the maximum feasible amount of idiosyncratic insurance, which is costless in equilibrium.

Rearranging Equation (22), expected excess returns can be expressed as

$$\mu^R - r = \underbrace{p^{ag}\sigma_A}_{\text{aggregate risk premium}} + \underbrace{p^{id}\phi\sigma_I}_{\text{idiosyncratic risk premium}}.$$

Hence, expected returns consist of compensation for both aggregate and idiosyncratic risk. In the absence of aggregate risk, growth, and adjustment costs, this reduces to $\alpha K^{\alpha-1} - r = p^{id}\phi\sigma_I$, so the marginal product of capital exceeds the interest rate whenever $\phi > 0$, even without borrowing constraints. Entrepreneurs refrain from expanding their business despite $\mu^R > r$ to limit risk exposure. In equilibrium, the wedge $\mu^R - r$ reflects risk aversion and the degree of diversification, as captured by ϕ.

Equation (23) characterizes demand for aggregate insurance. Entrepreneurs with high exposure to their business ($qk_{i,t} > n_{i,t}$) buy insurance, while those with low exposure ($qk_{i,t} < n_{i,t}$) provide it. In equilibrium, poorer entrepreneurs purchase aggregate insurance ($\theta^{ag}_{i,t} > 0$), while richer entrepreneurs supply it ($\theta^{ag}_{i,t} < 0$).

Finally, the consumption-wealth ratio in Equation (24) highlights three forces. First, the term $\bar{r}/(1 - \psi e^{-\bar{r}(T-a)})$ is the marginal propensity to consume (MPC), which rises with age as in finite-horizon models. Second, the bequest motive ψ shapes this profile: when $\psi = 0$, the MPC is constant, while when $\psi = 1$, wealth is fully consumed at T as in Merton (1969). Third, the consumption-wealth ratio scales with the human-financial wealth ratio, linking life-cycle consumption behavior to the empirical profiles documented in Section 2.

Taking stock. Proposition 1 yields three main lessons. First, entrepreneurs' risk-taking and consumption decisions are governed by the human-financial wealth ratio, $h_{i,t}/n_{i,t}$, which evolves systematically over the life cycle. Second, expected excess returns on entrepreneurial activity decompose into an aggregate and an idiosyncratic risk premium. Third, the idiosyncratic risk premium directly shapes the scale of entrepreneurial investment. Together, these mechanisms are central for the model's ability to replicate the empirical facts documented in Section 2.

4 Life-cycle patterns

We now explore the quantitative implications of the model for entrepreneurs' life-cycle behavior. We begin with the calibration and then turn to the life-cycle profiles for risk-taking, consumption, and financial wealth.

4.1 Risk-taking and savings over the life cycle

Technology, preferences, and demographics. Table 4 summarizes the calibration. The capital share is set to $\alpha = 0.33$ and depreciation to $\delta = 0.10$, standard in the literature (e.g., Campbell 1994). Productivity growth is $\mu = 0.003$, following Jeong and Townsend (2007) for Thailand. Investment follows $\Phi(\iota) = \sqrt{\overline{\Phi}^2 + 2\iota} - \overline{\Phi}$, the quadratic-adjustment-cost

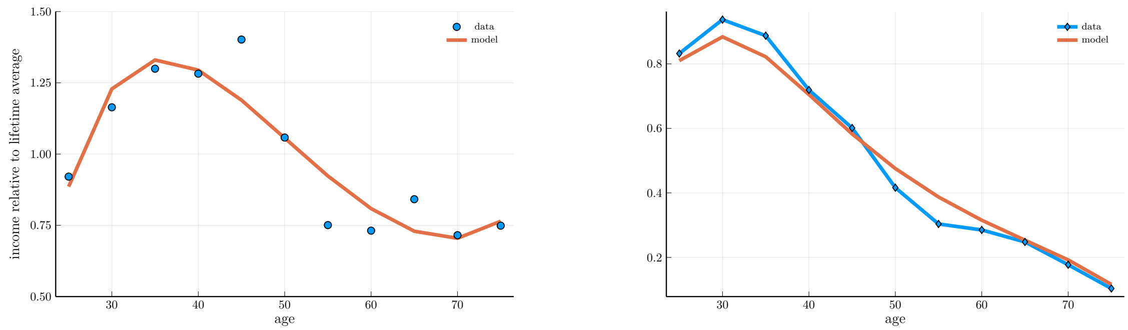
Table 4. Calibrated parameters

	Parameter	Choice
	Preferences	
ρ	Entrepreneur's rate of time preference	0.141
ρ_w	Wage earner's rate of time preference	0.127
ψ	Bequest motive	0.37
γ	Risk aversion	9.23
	Technology & financial friction	
μ_A	Average productivity growth rate	0.003
σ_A	Aggregate volatility	0.051
σ_{id}	Idiosyncratic volatility	0.203
ϕ	Moral hazard parameter	0.571
$\overline{\Phi}$	Adjustment cost parameter	0.90
α	Capital share in production function	0.33
δ	Depreciation rate	0.10
	Demographics	
g	Population growth	0.003
χ_e	Share of entrepreneurs in the population	0.46
T	Life span (adult life)	55

case, with $\overline{\Phi}$ chosen so that the long-run relative price of capital equals one. The wage earners' discount rate matches a real risk-free rate of $r = 3.5\%$, consistent with Thai data over the past 25 years. The entrepreneurs' discount rate and the bequest parameter are selected to match the consumption-wealth ratio at the beginning and end of life. We set the life horizon to $T = 55$ (ages 25-80) and population growth to $g = 0.3\%$, the most recent Thai value. Finally, χ_e is chosen to match the average share of business wealth in financial wealth. This implies that 46% of households are entrepreneurs, a fraction broadly consistent with the share of self-employed workers observed in Thailand.

Risk, return, and the moral hazard parameter. The volatility parameters, σ_A and σ_I, are chosen to match the aggregate and idiosyncratic components of total volatility. The aggregate risk premium is given by $p^{ag}\sigma_A = \gamma\sigma_A^2$, which allows us to pin down the coefficient of relative risk aversion γ from the observed aggregate risk premium. The estimated value of γ falls within the range considered plausible by Mehra and Prescott (1985) and

Figure 5: Life-cycle profiles: labor income and human-financial wealth ratio



(a) Labor income **(b)** Human-financial wealth ratio

is close to the benchmark values widely used in the asset pricing literature on the equity premium (see, e.g., Bansal and Yaron 2004).

The idiosyncratic risk premium is informative about the moral hazard parameter ϕ. When $\phi = 0$, entrepreneurs can fully insure against idiosyncratic shocks, so the idiosyncratic risk premium is zero. As ϕ increases, the share of non-diversifiable idiosyncratic risk rises, and so does the associated risk premium.

Using Equation (21) to solve for p^{id}, we can express the idiosyncratic risk premium in terms of the moral hazard parameter ϕ and observable quantities: the level of idiosyncratic volatility and the exposure of entrepreneurs' total wealth to their private business.[24] Given these quantities, we can identify ϕ.

Measuring human wealth. Following Gârleanu and Panageas (2015), we assume that labor supply varies deterministically with age according to

$$\bar{l}_{i,t} = \sum_{l=1}^{L} \Gamma_l e^{\varphi_l a_i},$$

with the normalization $\frac{\int_{\mathcal{E}_t} \bar{l}_{i,t}\, di}{\int_{\mathcal{E}_t} di} = 1$. This flexible functional form captures the empirical hump-shaped labor income profile while remaining analytically tractable. We estimate

[24]See Section 6 for a detailed discussion of the determination of p^{id} in equilibrium.

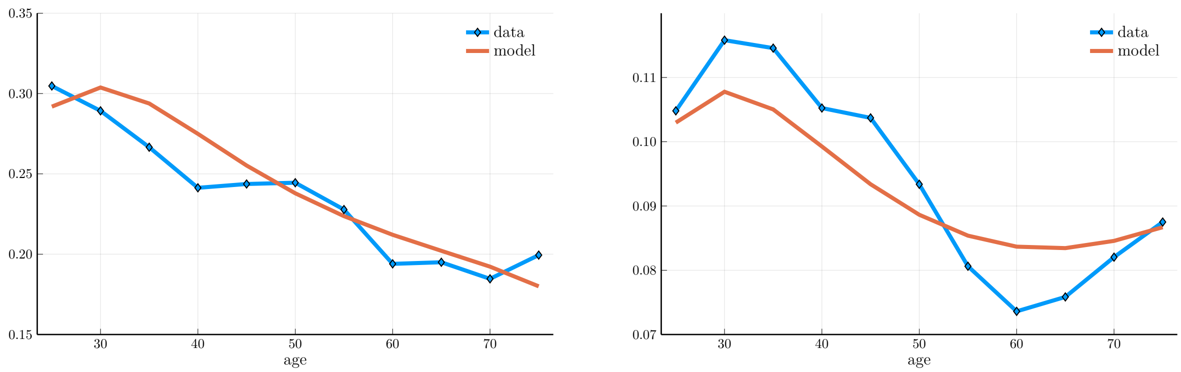
the parameters $(\Gamma_l, \varphi_l)_{l=1}^L$ by non-linear least squares, setting $L = 3$. The left panel of Figure 5 shows that the estimated profile closely matches the data.

We then consider the human-financial wealth ratio, $h_{i,t}/n_{i,t}$. For simplicity, we rely on the empirical age effects without additional controls. As shown in Appendix B, the life-cycle patterns remain virtually unchanged once year fixed effects or demographic controls are included. We focus on the more flexible specification with a larger number of age groups, as estimated in the appendix. The right panel of Figure 5 illustrates that the human-financial wealth ratio declines steadily with age. At the beginning of life, human wealth is nearly as large as financial wealth, but by age 50 it falls to roughly half of financial wealth. This decline reflects both relatively high labor income early in life and the progressively shorter horizon for future earnings.

Risk-taking and savings. The left panel of Figure 6 shows that our measure of risk-taking—the share of wealth invested in the business—declines with age, consistent with the evidence in Table 3. The model generates this pattern because effective risk aversion increases as the human-financial wealth ratio declines. Young entrepreneurs are therefore endogenously less risk-averse than older entrepreneurs, and they invest a larger fraction of their wealth in the business. Importantly, the ratio of risk-taking between the youngest and oldest entrepreneurs is pinned down entirely by the human-financial wealth ratio, which we calibrated independently of any information on the cross-section of entrepreneurial risk-taking.

The right panel of Figure 6 shows that the consumption-wealth ratio is U-shaped over the life cycle. This non-monotonic behavior reflects two opposing forces. First, the decline in the human-financial wealth ratio with age pushes entrepreneurs to reduce consumption. Second, the marginal propensity to consume increases with age, pushing them to consume more. The first force dominates early in life, while the second becomes stronger later in life, generating the U-shaped profile.

Figure 6: Life-cycle profiles: risk-taking and consumption-wealth ratio



(a) Share of wealth invested in the business

(b) Consumption-wealth ratio

4.2 Distributive implications of entrepreneurial risk

We now examine how entrepreneurial risk shapes wealth inequality. Our focus is on the joint distribution of (scaled) financial wealth and age, denoted by $f(n,a)$. From this joint distribution, we obtain, respectively, the average wealth by age and average entrepreneurial wealth:

$$n(a) = \int_{-h(a)}^{\infty} n f(n|a)\, dn, \qquad n_e = \int_0^T n(a) f(a)\, da,$$

where $f(a)$ is the age distribution.

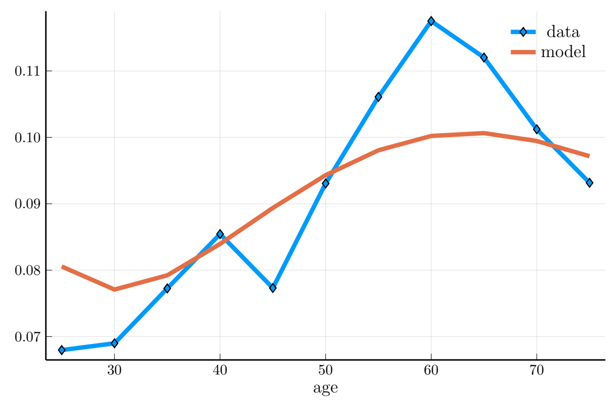
Between-group inequality. The next proposition characterizes average wealth by age.

Proposition 2. *Suppose the economy is in a stationary equilibrium. Then:*

i) Between-group inequality: *The wealth share of entrepreneurs of age a,* $\frac{f(a)n(a)}{n_e}$*, satisfies*

$$\log \frac{f(a)n(a)}{n_e} = \log \frac{f(0)n(0)}{n_e} + \underbrace{\log \left(\frac{1 + \frac{h(0)}{n(0)}}{1 + \frac{h(a)}{n(a)}} \right)}_{\substack{\text{human-to-financial} \\ \text{wealth effect}}} + \underbrace{\left[r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - (g + \mu_A) \right]}_{\text{generalized } r-g \text{ effect}} a - \underbrace{\int_0^a \frac{\bar{r}}{1 - \psi e^{-\bar{r}(T-a')}} da'}_{\text{average MPC effect}}. \tag{25}$$

Figure 7: Financial wealth distribution across age groups



where

$$n(0) = \left[e^{-\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-(g+\mu_A)-mpc_e\right)T} - 1\right]^{-1} h(0), \quad mpc_e = \frac{1}{T}\int_0^T \frac{\bar{r}}{1-\psi e^{-\bar{r}(T-a)}} da.$$

ii) Average financial wealth: *The average financial wealth of entrepreneurs is*

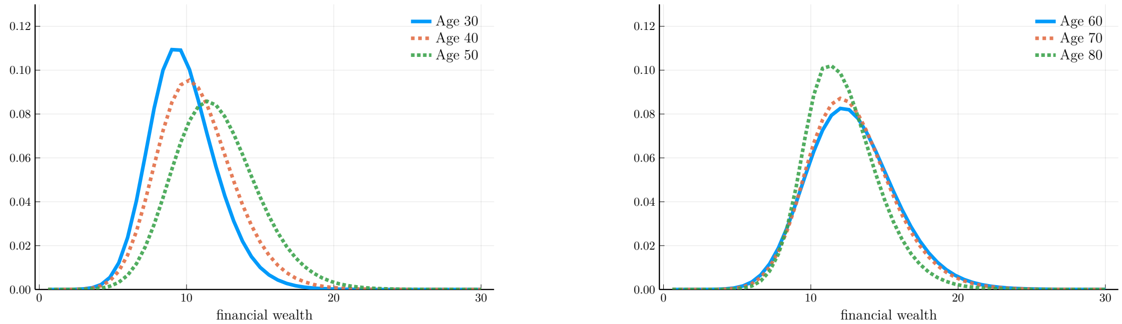
$$n_e = f(0)(n(0)+h(0)) \int_0^T e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-(g+\mu_A)\right)a} \frac{e^{-\bar{r}a} - \psi e^{-\bar{r}T}}{1-\psi e^{-\bar{r}T}} da - h_e, \tag{26}$$

where $h_e = \int_0^T f(a)h(a)\, da$.

Proposition 2 decomposes the wealth distribution across age groups into three effects: the *human-to-financial wealth effect*, as labor income accelerates the accumulation of financial wealth; the *generalized* $r-g$ *effect*, where the relevant return includes aggregate and idiosyncratic risk premia; and the *average MPC effect*, which captures how past consumption decisions shape current wealth. Part (ii) characterizes average entrepreneurial wealth, which determines the distribution of wealth between entrepreneurs and workers.

Figure 7 displays the share of wealth among entrepreneurs held by each age group. The model reproduces the inverted-U pattern of financial wealth observed in the data. For young entrepreneurs, the human-financial wealth effect and the generalized $r-g$

Figure 8: Stationary distribution of financial wealth by age



(a) Ages 30, 40, and 50

(b) Ages 60, 70, and 80

effect dominate the average MPC effect, so wealth shares initially rise with age. Later in life, the MPC effect takes over, bringing wealth shares down.

This hump-shaped pattern of wealth is a common implication of life-cycle models (see, e.g., Gomes 2020). Here, however, wealth accumulation responds to the price of idiosyncratic risk through the mechanisms highlighted above, which will be especially important in our counterfactual analysis of idiosyncratic insurance in Section 6.

Within-group inequality. We next characterize the distribution of wealth across entrepreneurs of the same age. Let $\mu_{n,t}(n,a)$ denote the drift (expected change) of financial wealth and $\sigma_{n,t}(n,a)$ its instantaneous volatility. The evolution of the conditional density $f(n|a)$ is governed by the Kolmogorov Forward Equation.

Lemma 2 (Kolmogorov Forward Equation)**.** *The conditional distribution of financial wealth $f_t(n|a)$ satisfies*

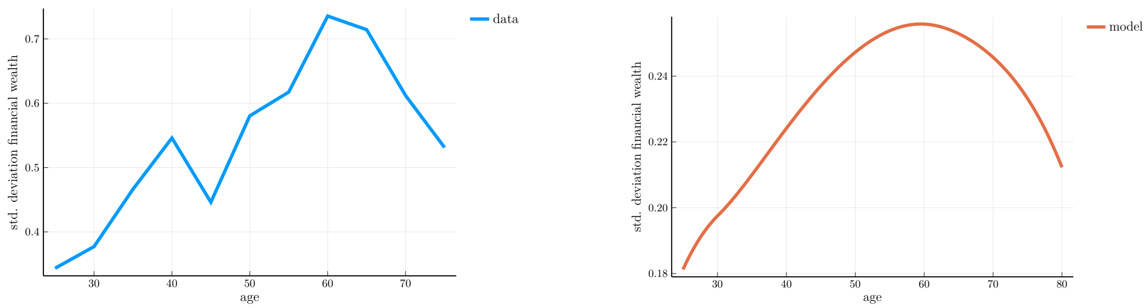
$$\frac{\partial f_t(n|a)}{\partial t}+\frac{\partial f_t(n|a)}{\partial a}=-\frac{\partial\left[f_t(n|a)\mu_{n,t}(n,a)\right]}{\partial n}+\frac{1}{2}\frac{\partial^2\left[f_t(n|a)\sigma^2_{n,t}(n,a)\right]}{\partial n^2}, \tag{27}$$

with boundary condition $f_t(e^{-gT}n|0)=f_t(n|T)$ and initial condition $f_0(n|a)$.

A closed-form solution can be obtained in the special case with no bequests ($\psi=1$).

Proposition 3 (Within-group inequality)**.** *Suppose $\psi=1$ and $r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}>\mu_A$.*

Figure 9: Standard deviation of financial wealth within age groups



(a) Within-group inequality: data

(b) Within-group inequality: model

Note: Panel (a) shows the standard deviation of financial wealth by age in the data. Panel (b) shows the standard deviation of financial wealth by age in the model economy. In both cases, we normalize the standard deviation by entrepreneurs' average wealth.

i) Shifted log-normal distribution. *Conditional on age, financial wealth follows a* shifted log-normal distribution *with support* $(-h(a), \infty)$*; equivalently,* $n + h(a)$ *is log-normal.*

ii) Variance by age. *The conditional variance of* n *is*

$$\mathbb{V}[n|a] = \left[e^{\left(\frac{p^{id}}{\gamma}\right)^2 a} - 1 \right] \left[h(0) e^{\left(r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \mu_A\right) a} \frac{e^{-\bar{r}a} - e^{-\bar{r}T}}{1 - e^{-\bar{r}T}} \right]^2. \tag{28}$$

iii) Inverted-U shape. *There exists* $0 < \hat{a} < T$ *such that* $\mathbb{V}[n|a]$ *increases with* a *for* $a < \hat{a}$ *and decreases for* $a > \hat{a}$.

Proposition 3 implies that wealth is *shifted log-normal*, with the shift $-h(a)$ reflecting the natural borrowing limit. Equation (28) describes how wealth dispersion evolves over the life cycle. Dispersion rises early in life, as entrepreneurs are hit by heterogeneous idiosyncratic shocks, and declines later as the increasing marginal propensity to consume counteracts the accumulation of risk.

The results in Proposition 3 assume no bequests. In the general case, we must use numerical methods. Figure 8 illustrates the stationary distribution of financial wealth for selected ages. The mean and the dispersion of the distribution initially increase with age, then both decline as entrepreneurs get older.

The same inverted U-pattern appears in the data, as shown in Figure 9. The standard deviation of n/n_e rises sharply until roughly age 60 and then falls by the end of the life cycle. Quantitatively, the model generates a substantial increase in inequality—by more than 40% from early in life until the peak—even though idiosyncratic return shocks are the sole source of heterogeneity across entrepreneurs. The increase in inequality is even more pronounced in the data. Allowing for differences in preferences or labor income could bring the within-group inequality in the model closer to the empirical evidence.

Consumption inequality and risk sharing. As shown in Appendix A.4, the variance of log total wealth—and thus log consumption—increases linearly with age: $\mathbb{V}[\log \omega_{i,t}|a] = \mathbb{V}[\log c_{i,t}|a] = \left(\frac{p^{id}}{\gamma}\right)^2 a$. The steepness of this life-cycle profile is determined by the extent of risk sharing in the economy, consistent with evidence that consumption dispersion can be used to infer the degree of incomplete markets (see, e.g., Storesletten et al. 2004). In our setting, the shadow price of idiosyncratic insurance p^{id} governs the slope, making consumption inequality a direct reflection of entrepreneurial risk exposure.

Extensions. Appendix D shows that our main findings remain robust when we incorporate three common features of the entrepreneurship literature: limited pledgeability of capital, endogenous occupational choice, and uninsurable labor income shocks. Taken together, these results show that a model with entrepreneurial risk and finite lives is able to generate realistic life-cycle profiles of risk-taking, consumption, wealth accumulation, and inequality.

5 Heterogeneous risk premia

The previous section showed that the baseline model reproduces the main life-cycle patterns of risk-taking, consumption, wealth accumulation, and inequality. We now turn to a second key empirical fact: the large heterogeneity in expected returns documented in Section 2. We show that an extension of the model can account for this feature as well.

5.1 Risk heterogeneity and decreasing returns to scale

Proposition 1 established that the expected return on the business includes compensation for idiosyncratic risk, consistent with the evidence in Table 2. However, in the baseline model the risk premium is equalized across entrepreneurs, at odds with the rich heterogeneity in expected returns observed in the data. This equalization holds even if entrepreneurs face different idiosyncratic volatilities: Equation (15) implies that the capital–labor ratio is the same across entrepreneurs, which through Equation (17) delivers the same expected return.

Introducing decreasing returns to scale. To generate heterogeneity in risk premia, the model must feature both heterogeneous exposure to idiosyncratic risk and *decreasing returns to scale* (DRS) in production. With constant returns to scale, entrepreneurs optimally choose the same capital–labor ratio, leading to identical expected returns regardless of their risk exposure. DRS breaks this equalization.

Formally, we extend the baseline technology by assuming that production depends not only on capital and labor but also on a fixed entrepreneurial ability e_i:

$$\tilde{y}_{i,t} = A_t k_{i,t}^{\alpha} l_{i,t}^{\beta} e_i^{1-\alpha-\beta}. \tag{29}$$

This span-of-control formulation follows Lucas (1978). For clarity, we first focus on the case where e_i is common across entrepreneurs and normalize $e_i = 1$. We also assume that entrepreneurs differ in their exposure to idiosyncratic risk, denoted by $\sigma_{I,i} \in \{\sigma_I^1, \ldots, \sigma_I^n\}$. To isolate the role of return heterogeneity, we abstract from overlapping generations by setting $g = 0$ and $T \to \infty$. The rest of the environment remains as described in Section 3.

Lemma 3 (Model with DRS)**.** *In a stationary equilibrium, the investment rate is given by* (16)*, the price of aggregate risk is $p^{ag} = \gamma\sigma_A$, and aggregate insurance is given by* (23)*. Labor demand* *is*

$$w_t = \beta k_{i,t}^{\alpha} l_{i,t}^{\beta-1},$$

and the conditional expectation of marginal returns is

$$\mu_{i,t}^{R} = \frac{\alpha \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} k_{i,t}^{\frac{\alpha+\beta-1}{1-\beta}} - \iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta. \tag{30}$$

Given value function $V_i(\omega)$, the optimality conditions for consumption and capital are

$$c_{i,t}^{-\gamma} = V_{i,\omega}\,, \qquad\qquad q k_{i,t} = -\frac{V_{i,\omega}}{V_{i,\omega\omega}} \frac{p_{i,t}^{id}}{\phi \sigma_{I,i}}.$$

Lemma 3 shows that with decreasing returns to scale, expected returns become *scale-dependent* whenever $\alpha + \beta < 1$. In this case, returns vary with the entrepreneur's business size, creating heterogeneity in premia. Importantly, heterogeneity arises through the dividend yield term, while expected capital gains remain common across entrepreneurs. This channel is consistent with how we measure entrepreneurial returns in the data, which primarily reflect differences in dividend yields.

Because scale-dependent returns break linear homogeneity, the model admits no closed-form solution. Policy functions are no longer linear in total wealth, and to make progress analytically we adapt perturbation methods in the spirit of Viceira (2001). We log-linearize policy functions around a risky steady state, yielding

$$\log c_j(\omega_{i,t}) = \log \bar{c}_j + \psi_{c,\omega}^{j} \hat{\omega}_{i,t} + \mathcal{O}(\hat{\omega}_{i,t}^2), \qquad \log k_j(\omega_{i,t}) = \log \bar{k}_j + \psi_{k,\omega}^{j} \hat{\omega}_{i,t} + \mathcal{O}(\hat{\omega}_{i,t}^2), \tag{31}$$

where $\hat{\omega}_{i,t} = \log \omega_{i,t} - \log \overline{\omega}_j$ and $\overline{\omega}_j = \exp \mathbb{E}[\log \omega_{i,t} \mid \sigma_{I,i} = \sigma_I^j]$. The approximation points $(\bar{c}_j, \bar{k}_j, \overline{\omega}_j)$ are endogenous, as in a risky steady-state approximation (see Coeurdacier et al., 2011). The following proposition delivers the main result of this section.

Proposition 4 (Heterogeneous risk premia)**.** *In the economy with decreasing returns to scale* *and heterogeneous idiosyncratic volatility, the price of idiosyncratic risk is*

$$p^{id}_{i,t} = \overline{p}^{id} - \psi^{j}_{p^{id},\omega}\hat{\omega}_{i,t}, \tag{32}$$

where $\mathbb{E}[p^{id}_{i,t}] = \overline{p}^{id}$ *is equalized across entrepreneurs, and* $\psi_{p^{id},\omega} > 0$*. Unconditional expected returns are then*

$$\mathbb{E}[\mu^{R}_{i,t}] = r + p^{ag}\sigma_A + \overline{p}^{id}\phi\sigma_{I,i}. \tag{33}$$

Proposition 4 establishes that risk premia are heterogeneous when entrepreneurs differ in their exposure to idiosyncratic volatility. Those operating riskier technologies earn higher expected returns in equilibrium, consistent with the evidence in Section 2. Equation (32) shows that conditional Sharpe ratios fluctuate with shocks to total wealth $\hat{\omega}_{i,t}$, but the unconditional Sharpe ratio is equalized across entrepreneurs. Hence, unconditional expected returns vary systematically with idiosyncratic volatility, generating heterogeneity in premia of the form observed in the data.

Skill heterogeneity. We abstracted from skill heterogeneity to focus on differences in risk. Consider now the polar opposite assumption: $e_i \in \{e^1, \ldots, e^n\}$ and $\sigma_{I,i} = \sigma_I$.

Proposition 5 (Skill heterogeneity)**.** *Capital allocations satisfy* $\overline{k}_j = e_j K$*. The price of idiosyncratic risk remains given by Equation* (32)*, implying* $\mathbb{E}[p^{id}_{i,t}] = \overline{p}^{id}$ *since* $\sigma_{I,i} = \sigma_I$*. Hence, the unconditional expectation of returns is equalized across entrepreneurs.*

Proposition 5 shows that differences in skill alone cannot explain the heterogeneity in risk premia observed in the data. More skilled entrepreneurs operate larger businesses, but their expected returns do not exceed those of less skilled entrepreneurs. This logic is closely related to Berk and Green (2004), who argue that in mutual funds greater managerial skill translates into larger fund size rather than higher average returns. In our setting with decreasing returns to scale, ability similarly scales up business size, while equilibrium expected returns remain pinned down by risk exposure, as in Equation (33).

5.2 The role of borrowing constraints and differences in net worth

We have seen that heterogeneity in expected returns arises entirely from differences in risk exposure. This result, however, relies on the absence of borrowing constraints. Suppose instead that pledgeability of physical capital is limited, so entrepreneurs face collateral constraints.[25] In that case, expected returns should vary with net worth: constrained entrepreneurs with low wealth face higher returns, everything else equal. Thus, a central prediction of a collateral-constraint model is that variations in financial wealth drive differences in expected returns, once we control for risk exposure.

Could differences in net worth explain the heterogeneity in expected returns described in Section 2? We test this prediction in the data. The left panel of Figure 10 plots average returns against average net worth across entrepreneurs. The relationship is statistically insignificant and, if anything, slightly positive. The right panel refines the test by focusing on entrepreneurs with below-average net worth—those most likely to be constrained—and orthogonalizing wealth with respect to aggregate beta and idiosyncratic volatility. Again, the association between net worth and returns is statistically insignificant, with net worth explaining less than half a percentage point of the variation in returns.[26]

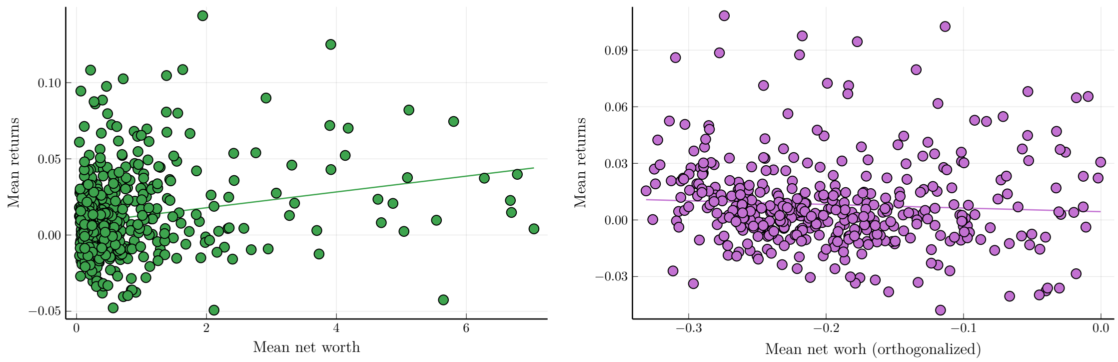
Moreover, a collateral-constraint model predicts that expected returns should be unrelated to risk exposure among constrained entrepreneurs. The strong risk–return relationship in Section 2 therefore suggests collateral constraints are not the main driver of returns. Testing this more directly, we repeat the regressions for below-average-wealth entrepreneurs and find an even stronger risk–return relationship: $R^2 = 0.82$ versus 0.68 in the full sample.

The lack of a systematic link between net worth and expected returns is consistent with Proposition 4, which predicts that wealth does not matter once risk exposure is con-

[25] See Appendix D.1 for details.

[26] Entrepreneurs with low average net worth are more likely to be constrained in models with transitory differences in skill. This may not hold with permanent skill differences. Empirically, we find no systematic relationship between expected returns and net worth among high net-worth entrepreneurs either.

Figure 10: Average returns vs. average net worth



Note: The left panel plots average net worth against average returns by entrepreneur, with net worth normalized by its cross-sectional mean. The right panel plots the residual component of net worth (orthogonal to aggregate beta and idiosyncratic variance) against average returns for entrepreneurs with below-average orthogonalized net worth. To limit the influence of outliers, we trim 1% of the observations in both tails.

trolled for.[27] We conclude that collateral constraints—and differences in net worth more broadly—are not the primary drivers of expected entrepreneurial returns in our data.

6 Counterfactuals

Having shown that the model captures the key patterns in the data, we now turn to counterfactual experiments. We study how changes in idiosyncratic insurance affect aggregate outcomes. Financial development, captured by the moral hazard parameter ϕ, is closely tied to economic development. We first analyze the long-run effects of a reform that improves insurance and then examine its transitional dynamics.

6.1 Aggregate capital and idiosyncratic risk premium

We now characterize the determination of aggregate variables in a stationary equilibrium. Because the model always features decreasing returns to scale at the aggregate level, we revert to the baseline case with constant returns to scale at the individual level and homogeneous idiosyncratic volatility, $\sigma_{I,i} = \sigma_I$. Allowing for decreasing returns at

[27]For example, under skill heterogeneity, more skilled entrepreneurs accumulate greater wealth but earn the same expected return as less skilled entrepreneurs with the same risk exposure.

the individual level does not materially affect the aggregate implications.

The interest rate is given by the standard condition

$$r = \rho_w + \gamma\mu_A - \tfrac{\gamma(\gamma+1)}{2}\sigma_A^2,$$

as derived in Appendix C.1. From Equation (16), we obtain $q = \overline{\Phi} + g + \delta$, and we choose $\overline{\Phi}$ such that $q = 1$. From Proposition 1, the price of aggregate risk is $p^{ag} = \gamma\sigma_A$.

Equation (22) implies

$$r + p^{ag}\sigma_A + p^{id}\phi\sigma_I = \frac{\alpha K^{\alpha-1} - \iota(q)}{q} + g + \mu_A. \tag{34}$$

The left-hand side gives the required rate of return to invest in the business, while the right-hand side gives the actual expected return, determined by the marginal product of capital (MPK). Equation (34) generalizes the standard textbook relation between the MPK and the interest rate, since in the absence of growth and risk we recover $r = \alpha K^{\alpha-1} - \delta$.

Expression (34) implies an inverse relation between the idiosyncratic risk premium $p^{id}\phi\sigma_I$ and the capital–labor ratio K. This downward-sloping relationship, shown by the solid blue line in the left panel of Figure 11, is the *MPK schedule*. It is analogous to the cross-section relationship derived in the DRS model of Section 5, since at the aggregate level production features DRS due to the fixed labor supply.

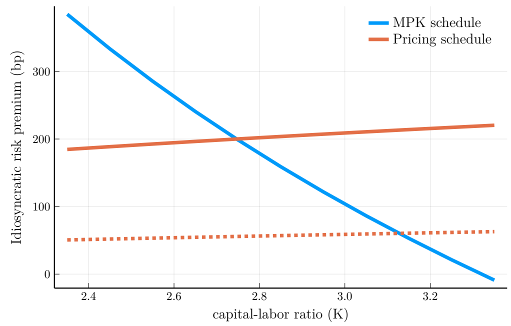
To close the system, we need an additional condition linking K and p^{id}. Aggregating the capital demand condition (21) across entrepreneurs yields

$$p^{id} = \underbrace{\gamma}_{\text{risk aversion}} \underbrace{\phi\sigma_I}_{\text{effective risk}} \underbrace{\frac{qK}{\chi_e(n_e + h_e)}}_{\text{id. risk exposure}}, \tag{35}$$

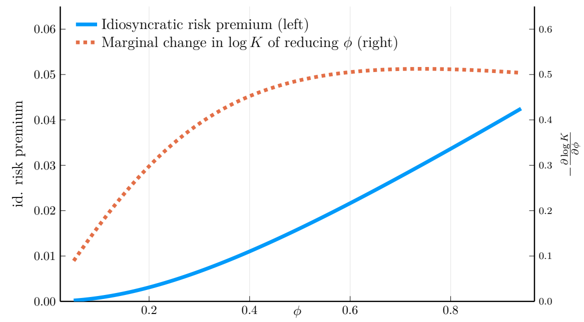
where n_e and h_e denote the average financial and human wealth of entrepreneurs.

The price of idiosyncratic risk depends on risk aversion γ and the effective level of

Figure 11: Idiosyncratic Risk Premium and Capital Stock



(a) MPK schedule vs. Pricing schedule



(b) Idiosyncratic risk premium and $-\frac{\partial \log K}{\partial \phi}$

Note: In the left panel, the solid (dashed) upward-sloping curve shows the pricing schedule in the initial (new) stationary equilibrium for $\phi_0 = 0.571$ ($\phi_1 = 0.286$). The right panel shows the equilibrium idiosyncratic risk premium and the marginal increase in the capital stock by reducing the moral hazard parameter for different values of ϕ.

idiosyncratic risk, $\phi\sigma_I$. It also depends on the *idiosyncratic risk exposure*, i.e. the ratio of physical assets to the total wealth of entrepreneurs. When entrepreneurs are more exposed to their businesses, they require a higher idiosyncratic risk premium.

Substituting Equation (26) to eliminate $n_e + h_e$, we obtain an implicit relationship between p^{id} and K. The left panel of Figure 11 depicts this relationship as the solid upward-sloping curve, which we refer to as the *pricing schedule*. The intersection of the MPK schedule and the pricing schedule jointly determines the equilibrium values of p^{id} and K.

Idiosyncratic Sharpe ratio. The model matches both the risk premium and the volatility of aggregate and idiosyncratic risk reported in Table 2. A striking fact is that the Sharpe ratio for aggregate risk is nearly five times larger than the one for idiosyncratic risk, even though most of the risk is idiosyncratic. Expression (35) helps explain this pattern.

The Sharpe ratio of aggregate risk is $p^{ag} = \gamma\sigma_A$. By contrast, the Sharpe ratio of idiosyncratic risk is $p^{id}\phi$, since the reported volatility does not account for insurance. If idiosyncratic risk were priced analogously to aggregate risk, the implied Sharpe ratio would be $\gamma\sigma_I$, over four times larger than the aggregate counterpart.

However, from the pricing equation we obtain

$$p^{id}\phi = \gamma\sigma_I\phi^2 \frac{qK/\chi_e}{n_e + h_e}.$$

The moral hazard parameter $\phi = 0.57$ reduces $p^{id}\phi$ by $1 - \phi^2 \approx 67\%$. The remainder of the adjustment comes from the risk-exposure factor, with $\frac{qK}{\chi_e(n_e+h_e)} \approx 0.16$. Thus, the low price of idiosyncratic risk arises either because entrepreneurs are partially insured or because only a fraction of their total wealth is directly exposed to business risk.

Importantly, without accounting for human wealth and heterogeneous agents, one would wrongly attribute the low Sharpe ratio entirely to high levels of insurance.

6.2 Long-run effects of relaxing insurance constraints

We now turn to the aggregate implications of relaxing insurance constraints. High values of ϕ capture situations where access to insurance arrangements—formal or informal—is limited. As institutional arrangements improve, such frictions decline and entrepreneurs bear less risk. Financial development is thus represented in the model as a reduction in the moral hazard parameter ϕ.

Interpretation. How should we interpret changes in the moral hazard parameter ϕ? One natural interpretation, consistent with the market structure in Section 3, is an expansion of insurance available to entrepreneurs. Karlan et al. (2014) study such an intervention in northern Ghana, where entrepreneurs were offered rainfall insurance. They find strong demand for insurance and substantial increases in investment and risk taking, consistent with limited insurance being the binding constraint.

Other interventions may also map into our framework. Battaglia et al. (2024) study a *flexible lending contract* in Bangladesh that allows entrepreneurs to postpone repayments after adverse shocks. They find that such contracts raise investment and risk taking, again consistent with insurance provision being the relevant mechanism.[28] Appendix E.3 shows that the optimal allocation can be implemented through this type of contract: in contrast to a standard, non–state-contingent loan, a flexible contract embeds partial insurance by making repayments contingent on realized shocks. In this setting, the skin-in-the-game constraint (11) is replaced by a credit limit that ultimately governs the degree of insurance. A reduction in ϕ can thus be interpreted as an expansion of this limit—an improvement in state-contingent credit rather than a pure increase in borrowing capacity.

Overall, the formulation is flexible enough to encompass a range of institutional arrangements that expand insurance to entrepreneurs, whether through formal insurance markets or flexible financial contracts.

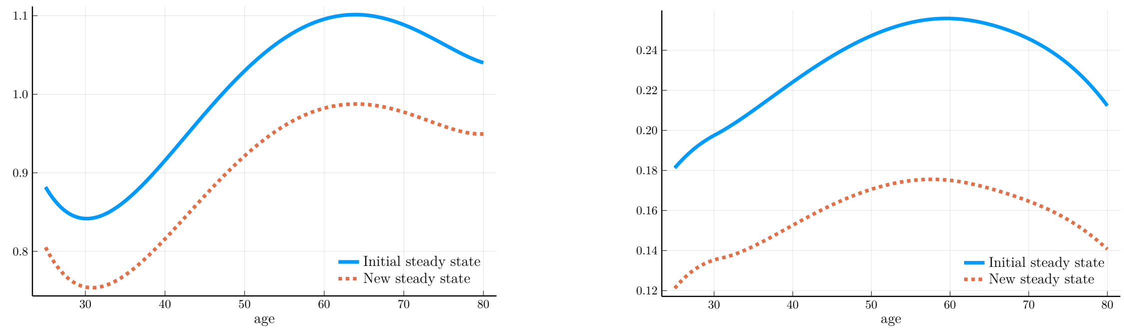
Long-run effects. Panel (a) of Figure 11 shows the impact of reducing ϕ. The MPK schedule is unchanged, but the pricing schedule shifts downward. In the long run, lower expected returns reduce the MPK and raise the capital stock. For instance, halving the moral hazard parameter from $\phi_0 = 0.57$ to $\phi_1 = 0.29$ lowers the idiosyncratic risk premium by about 140 basis points and increases the capital stock by roughly 13%.

Panel (b) of Figure 11 highlights that the economy's response to changes in ϕ is highly non-linear. The idiosyncratic risk premium (solid line) is convex in ϕ, so reductions in the moral hazard parameter have stronger effects in economies with limited financial development. The dashed line shows the marginal increase in capital due to a reduction in ϕ: the same decline in ϕ raises the capital stock by a factor five more in a low-financial-development economy (high ϕ) than in a high-financial-development economy (low ϕ).

Financial development also has important implications for inequality. The left panel of Figure 12 shows the average financial wealth for each age group relative to average entrepreneurial wealth on the initial stationary equilibrium. Financial wealth falls for all age groups in the new equilibrium, as it is harder to accumulate wealth with lower

[28] Cordaro et al. (2022) find positive effects on investment and profits from a microcredit contracts with performance-contingent repayment. A related literature studies the role of flexible repayment in microcredit, see, e.g., Aragón et al. (2020) and Barboni and Agarwal (2023).

Figure 12: Financial development and inequality in the long run



(a) Between-group inequality **(b)** Within-group inequality

Note: Inequality in a stationary equilibrium for $\phi_0 = 0.571$ (initial steady state) and $\phi_1 = 0.286$ (new steady state). Panel (a) shows average financial wealth by age. Panel (b) shows the standard deviation of financial wealth by age. All variables are normalized by entrepreneurs' average wealth in the initial steady state.

expected returns. In the long run, inequality falls after a reduction in ϕ, as shown in the right panel of Figure 12, given that entrepreneurs are less exposed to risk.

6.3 Dynamic effects of relaxing risk constraints

So far, we have compared two stationary equilibria. To compute the welfare implications of relaxing risk constraints, however, it is essential to account for what happens during the transition.[29] We now turn to the dynamic effects of the reform.

Computing the transitional dynamics. We consider a "small open economy" version of the model, in which the interest rate stays fixed at the level of the original stationary equilibrium.[30] This assumption allows us to isolate the role of fluctuations in the idiosyncratic risk premium. The next proposition characterizes the evolution of aggregate variables during the transition.

[29] While equilibrium variables remain exposed to aggregate shocks, we compute the transitional dynamics for *scaled variables*. For example, we track $q_t \equiv \tilde{q}_t / A_t$, where $\tilde{q}_t$ moves with the realization of A_t and the deterministic path of q_t.

[30] Appendix D.4 shows that in a closed economy the interest rate also remains constant during the transition if wage earners have Epstein–Zin utility with linear intertemporal preferences.

Proposition 6. *The price of aggregate risk is* $p_t^{ag} = \gamma \sigma_A$. *The evolution of* $(q_t, K_t, \{\zeta_t(a), h_t(a), \omega_t(a)\})$ *is given by a pair of ODEs:*

$$
\begin{aligned}
\dot{K}_t &= \left[\Phi(\iota(q_t)) - \delta - g\right] K_t, \\
\dot{q}_t &= \left[r + \gamma \sigma_A^2 + \gamma \phi^2 \sigma_I^2 \frac{q_t K_t}{\chi_e \omega_{e,t}} + \delta - \mu_A - \Phi(\iota(q_t)) - \frac{\alpha K_t^{\alpha - 1} - \iota(q_t)}{q_t} \right] q_t,
\end{aligned}
$$

and three PDEs:

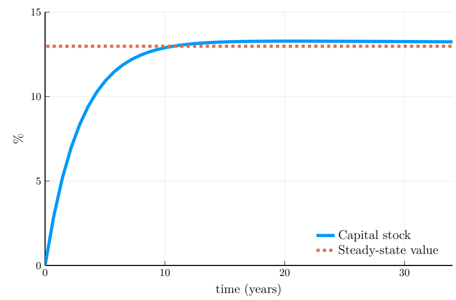
$$
\begin{aligned}
\frac{\partial \zeta_t(a)}{\partial t} &= -\frac{\partial \zeta_t(a)}{\partial a} + \zeta_t^2(a) - \bar{r}_t \zeta_t(a), \\
\frac{\partial h_t(a)}{\partial t} &= -\frac{\partial h_t(a)}{\partial a} + (r + \gamma \sigma_A^2 - \mu_A) h_t(a) - (1 - \alpha) K_t^{\alpha} \bar{l}(a), \\
\frac{\partial \omega_t(a)}{\partial t} &= -\frac{\partial \omega_t(a)}{\partial a} + \left[r + \gamma \sigma_A^2 - \mu_A + \gamma \phi^2 \sigma_I^2 \left(\frac{q_t K_t}{\chi_e \omega_{e,t}} \right)^2 - \zeta_t(a) \right] \omega_t(a),
\end{aligned}
$$

subject to the boundary conditions described in the appendix.

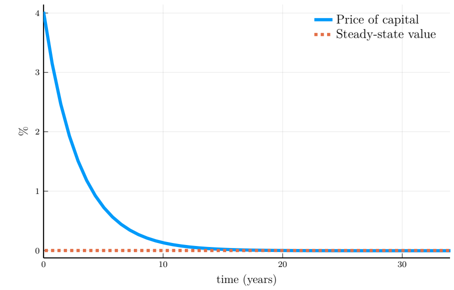
The first ODE aggregates (7), while the second is the time-dependent counterpart of (34). The PDE for the consumption–wealth ratio $\zeta_t(a)$ comes from the HJB equation; the PDE for $h_t(a)$ is the dynamic version of (18); and the PDE for $\omega_t(a)$ follows from averaging entrepreneurs' budget constraints.

Transitional dynamics in heterogeneous-agent models are often computed using shooting algorithms (see, e.g., Guerrieri and Lorenzoni 2017, Achdou et al. 2017). As such methods are impractical here, we adopt a hybrid approach combining perturbation and finite differences. We first discretize the system of ODE/PDEs using finite differences, then linearize around the new stationary equilibrium. Unlike approaches that linearize around a shock-free economy (e.g., Ahn et al. 2018), we allow for aggregate shocks. Variables in levels remain proportional to A_t (e.g., $\tilde{q}_t = q_t A_t$). This method avoids small-risk approximations and captures time-varying risk premia and precautionary savings effects.

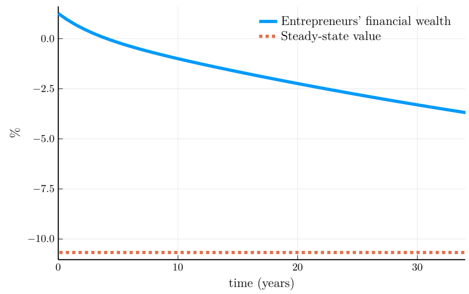
Figure 13: Transitional dynamics: aggregate variables



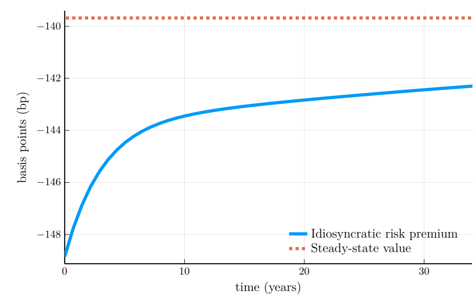
(a) Capital Stock



(b) Price of Capital



(c) Entrepreneurs' Financial Wealth

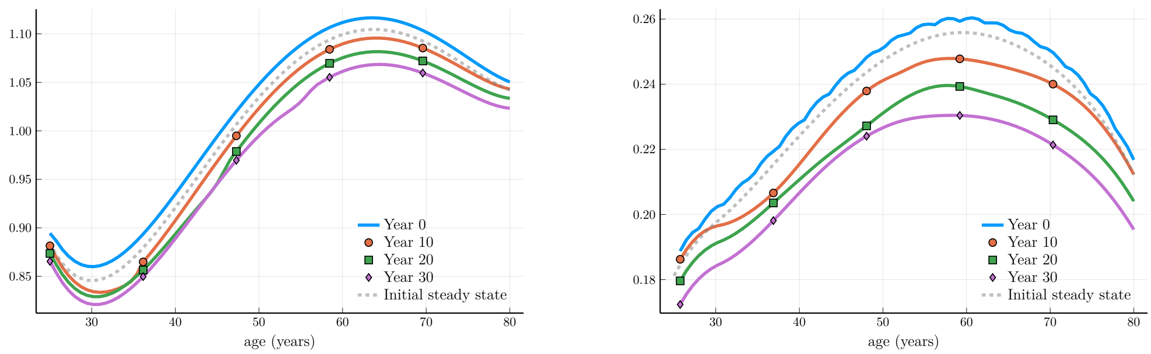


(d) Idiosyncratic Risk Premium

Note: Transitional dynamics from a stationary equilibrium with $\phi_0 = 0.571$ to a stationary equilibrium with $\phi_1 = 0.286$. Capital stock, the relative price of capital, and entrepreneurs' financial wealth are expressed as percentage deviations from the initial steady state. The idiosyncratic risk premium is expressed as absolute deviation from the initial steady state in basis points.

Short-run dynamics: the overshooting effect. Figure 13 displays the transitional dynamics for aggregate variables. We consider a reform that reduces the moral hazard parameter by half, from $\phi_0 = 0.571$ under our calibration to $\phi_1 = 0.286$. The reform triggers an investment boom and a sharp rise in business valuations that lasts for roughly a decade, as shown in Panels (a) and (b). The short-run response of q exceeds its long-run level by a wide margin, generating an *overshooting* effect. As entrepreneurs bear less risk, they demand a smaller premium. In the long run, this leads to a larger capital stock. But since capital is fixed in the short run, expected returns can only fall through expected capital losses. Hence the price of capital jumps on impact and then gradually converges back to its new steady-state level. This logic is reminiscent of Dornbusch (1976), where exchange rates react more strongly in the short run to generate the capital losses required to restore equilibrium. The overshooting mechanism has important consequences for

Figure 14: Transitional dynamics: inequality



(a) Between-group inequality

(b) Within-group inequality

Note: Transitional dynamics from a stationary equilibrium with $\phi_0 = 0.571$ to a stationary equilibrium with $\phi_1 = 0.286$. Panel (a) shows average financial wealth by age. Panel (b) shows the standard deviation of financial wealth by age.

wealth dynamics and inequality.

While entrepreneurs' wealth declines in the long run, it increases in the short run due to a revaluation effect. Financial wealth jumps up on impact, but lower expected returns slow down subsequent accumulation. This effect unfolds gradually, as part of the adjustment operates through reduced bequests. Even thirty years after the reform, the reduction in financial wealth is only about 25% of its eventual long-run decline.

Finally, Panel (d) shows that the idiosyncratic risk premium also overshoots: its short-run response exceeds the new steady-state level in magnitude by nearly ten basis points. After a decade, more than half of the gap to the steady state is closed, and the premium then rises gradually as entrepreneurs' wealth erodes.

Kuznets dynamics. The left panel of Figure 14 plots average financial wealth by age at different points in time, normalized by average entrepreneurial wealth in the initial equilibrium. Because of the revaluation effect, financial wealth rises immediately on impact. The effect is stronger for younger entrepreneurs, who are more exposed to the business. Over time, financial wealth declines as lower expected returns gradually dominate.

The right panel of Figure 14 shows the standard deviation of financial wealth by age.

Here, too, the short-run and long-run responses differ. Wealth inequality rises in the short run but falls in the long run. Wealthier entrepreneurs benefit disproportionately from the reform, since they hold more capital.

Demographics further shape these dynamics. After the initial increase, wealth dispersion eventually declines for all age groups. Yet entrepreneurs entering their careers just after the reform are most affected. Ten years after the reform, inequality has fallen by twice as much for 35-year-old entrepreneurs, who spent their entire careers under the new regime, compared to 80-year-olds, who lived most of their working lives under the old regime. Similarly, the decline in inequality is more pronounced for 35-year-olds than for 25-year-olds.

Taking stock. The reform initially generates an investment boom and a temporary rise in wealth inequality. The surge in investment and risk taking is consistent with the experimental evidence of Karlan et al. (2014) and Battaglia et al. (2024). As the economy converges to its new steady state, inequality declines, settling at a lower level in the long run. In other words, inequality rises during the high-growth phase and falls once the economy reaches a higher level of development, consistent with Kuznets's (1955) curve.[31]

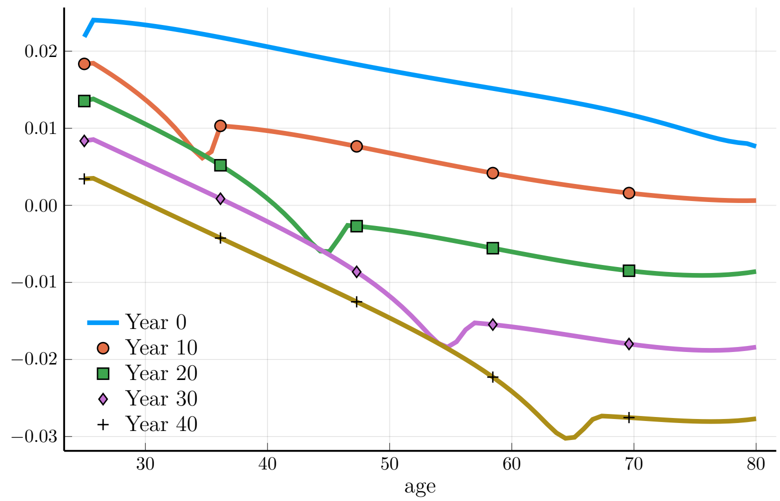
Welfare implications. We now turn to the welfare implications of insurance constraints. Recall that the value function of an entrepreneur of age a at time t is

$$V_t(n,a) = \zeta_t(a)^{-\gamma}\frac{(n+h_t(a))^{1-\gamma}}{1-\gamma}.$$

Welfare depends on financial wealth n, human wealth $h_t(a)$, and the consumption–wealth ratio $\zeta_t(a)$, which reflects the path of expected future returns. We evaluate financial wealth at the average level of each age group, $n = n_t(a)$, and apply a monotonic trans-

[31] Moll (2012) derive a Kuznets curve by showing that the steady-state top wealth share is hump-shaped in financial development. In contrast, we emphasize transitional dynamics rather than long-run steady-state comparisons.

Figure 15: Transitional dynamics: welfare



formation of the value function to measure welfare in consumption units. Hence, our welfare metric is

$$\mathcal{W}_t(a) = \log\left[u^{-1}(V_t(n_t(a), a))\right] - \log\left[u^{-1}(V^*(n^*(a), a))\right] = \frac{\gamma \hat{\zeta}_t(a)}{\gamma - 1} + \hat{\omega}_t(a),$$

where $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ and a hat denotes log deviations from the initial steady state.

Figure 15 plots welfare gains by age group over time. The generation alive at the time of the reform benefits the most, especially younger entrepreneurs, due to the revaluation effect. By contrast, future generations experience welfare losses: lower expected returns reduce wealth accumulation and bequests, which diminishes welfare over time. Ten years after the reform, entrepreneurs entering the labor force (ages 25–35) have smaller welfare gains than their counterparts of the same age at the time of the reform. Thirty years after the reform, entrepreneurs aged 35–55 are worse off relative to the no-reform counterfactual, with the largest losses accruing to older entrepreneurs. Thus, the initial generation of entrepreneurs reaps most of the welfare benefits of financial development, while wage earners also benefit as wages rise with higher capital accumulation in the economy.

7 Conclusion

We study the aggregate and distributive implications of entrepreneurial risk through a life-cycle model with aggregate and idiosyncratic shocks under limited insurance. The model quantitatively replicates key empirical patterns: risk-taking and savings profiles over the life cycle, the inverted-U shape of wealth inequality, and the magnitude of aggregate and idiosyncratic risk premia. An extension with decreasing returns to scale and heterogeneous volatility further accounts for the cross-sectional heterogeneity in risk premia documented in the data.

Our results underscore the central role of limited risk sharing in shaping both the pace of economic development and the evolution of wealth inequality. They also point to the broad implications of financial innovations and policy reforms that relax insurance frictions—policies that affect not only aggregate capital accumulation, but also the distribution of wealth and welfare across generations.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and wealth distribution in macroeconomics: A continuous-time approach," *The Review of Economic Studies*, 2017, *84* (1), 1–45.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf, "When inequality matters for macro and macro matters for inequality," *NBER Macroeconomics Annual*, 2018, *32* (1), 1–75.

Aiyagari, S Rao, "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Angeletos, George-Marios, "Uninsured idiosyncratic investment risk and aggregate saving," *Review of Economic Dynamics*, 2007, *10* (1), 1–30.

Aragón, Fernando M., Alexander Karaivanov, and Karuna Krishnaswamy, "Credit Lines in Microcredit: Short-Term Evidence from a Randomized Controlled Trial in India," *Journal of Development Economics*, 2020, *146*, 102497.

Bach, Laurent, Laurent E Calvet, and Paolo Sodini, "Rich pickings? Risk, return, and skill in household wealth," *American Economic Review*, 2020, *110* (9), 2703–47.

Bansal, Ravi and Amir Yaron, "Risks for the long run: A potential resolution of asset pricing puzzles," *The journal of Finance*, 2004, *59* (4), 1481–1509.

Barboni, Giorgia and Parul Agarwal, "How do flexible microfinance contracts improve repayment rates and business outcomes? Experimental evidence from India," *Experimental Evidence from India (February 14, 2023)*, 2023.

Barro, Robert J, "Rare disasters, asset prices, and welfare costs," *American Economic Review*, 2009, *99* (1), 243–64.

Battaglia, Marianna, Selim Gulesci, and Andreas Madestam, "Repayment Flexibility and Risk Taking: Experimental Evidence from Credit Contracts," *Review of Economic Studies*, 2024, *91* (5), 2635–2675.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S Goldstein, "Portfolio choice over the life-cycle when the stock and labor markets are cointegrated," *The Journal of Finance*, 2007, *62* (5), 2123–2167.

Berk, Jonathan B and Richard C Green, "Mutual fund flows and performance in rational markets," *Journal of political economy*, 2004, *112* (6), 1269–1295.

Black, Fischer, Michael C Jensen, and Myron S Scholes, "The capital asset pricing model: Some empirical tests," 1972.

Boar, Corina, Denis Gorea, and Virgiliu Midrigan, "Why Are Returns to Private Business Wealth So Dispersed?," Technical Report, National Bureau of Economic Research 2022.

Bodie, Zvi, Robert C Merton, and William F Samuelson, "Labor supply flexibility and portfolio choice in a life cycle model," *Journal of Economic Dynamics and Control*, 1992, *16* (3), 427–449.

Buera, Francisco J and Yongseok Shin, "Financial frictions and the persistence of history: A quantitative exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272.

Burmeister, Edwin and Marjorie B McElroy, "Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory," *The Journal of Finance*, 1988, *43* (3), 721–733.

Campbell, John Y, "Inspecting the mechanism: An analytical approach to the stochastic growth model," *Journal of Monetary Economics*, 1994, *33* (3), 463–506.

Chamberlain, Gary and Charles A Wilson, "Optimal intertemporal consumption under uncertainty," *Review of Economic dynamics*, 2000, *3* (3), 365–395.

Chen, Hui, Jianjun Miao, and Neng Wang, "Entrepreneurial finance and nondiversifiable risk," *The Review of Financial Studies*, 2010, *23* (12), 4348–4388.

Chen, Zhiyao and Ilya A Strebulaev, "Macroeconomic risk and idiosyncratic risk-taking," *The Review of Financial Studies*, 2019, *32* (3), 1148–1187.

Christiano, Lawrence J, Roberto Motto, and Massimo Rostagno, "Risk shocks," *American Economic Review*, 2014, *104* (1), 27–65.

Cochrane, John H, *Asset pricing: Revised edition*, Princeton university press, 2009.

Coeurdacier, Nicolas, Helene Rey, and Pablo Winant, "The risky steady state," *American Economic Review*, 2011, *101* (3), 398–401.

Constantinides, George M and Darrell Duffie, "Asset pricing with heterogeneous consumers," *Journal of Political economy*, 1996, *104* (2), 219–240.

Cordaro, Francesco, Marcel Fafchamps, Colin Mayer, Muhammad Meki, Simon Quinn, and Kate Roll, "Microequity and Mutuality: Experimental Evidence on Credit with Performance-Contingent Repayment," Working Paper 30411, National Bureau of Economic Research September 2022.

David, Joel M, Lukas Schmid, and David Zeke, "Risk-adjusted capital allocation and misallocation," *Journal of Financial Economics*, 2022, *145* (3), 684–705.

DeMarzo, Peter M and Michael J Fishman, "Optimal long-term financial contracting," *The Review of Financial Studies*, 2007, *20* (6), 2079–2128.

_ **and Yuliy Sannikov**, "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *The Journal of Finance*, 2006, *61* (6), 2681–2724.

Detemple, Jerome and Shashidhar Murthy, "Equilibrium asset prices and no-arbitrage with portfolio constraints," *The Review of Financial Studies*, 1997, *10* (4), 1133–1174.

Di Tella, Sebastian, "Uncertainty shocks and balance sheet recessions," *Journal of Political Economy*, 2017, *125* (6), 2038–2081.

Doepke, Matthias and Michèle Tertilt, "Families in macroeconomics," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 1789–1891.

Dornbusch, Rudiger, "Expectations and exchange rate dynamics," *Journal of political Economy*, 1976, *84* (6), 1161–1176.

Fagereng, Andreas, Martin Blomhoff Holm, Benjamin Moll, and Gisle Natvik, "Saving behavior across the wealth distribution: The importance of capital gains," Technical Report, National Bureau of Economic Research 2019.

Fama, Eugene F and James D MacBeth, "Risk, return, and equilibrium: Empirical tests," *Journal of political economy*, 1973, *81* (3), 607–636.

_ and Kenneth R French, "The cross-section of expected stock returns," *the Journal of Finance*, 1992, *47* (2), 427–465.

Gârleanu, Nicolae and Stavros Panageas, "Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing," *Journal of Political Economy*, 2015, *123* (3), 670–685.

_ , _ , and Jianfeng Yu, "Financial entanglement: A theory of incomplete integration, leverage, crashes, and contagion," *American Economic Review*, 2015, *105* (7), 1979–2010.

Giglio, Stefano and Dacheng Xiu, "Asset pricing with omitted factors," *Journal of Political Economy*, 2021, *129* (7), 1947–1990.

_ , Yuan Liao, and Dacheng Xiu, "Thousands of alpha tests," *The Review of Financial Studies*, 2021, *34* (7), 3456–3496.

Gomes, Francisco, "Portfolio choice over the life cycle: A survey," *Annual Review of Financial Economics*, 2020, *12* (1), 277–304.

Gomez, Matthieu, "Asset prices and wealth inequality," 2017.

Grossman, Sanford J and Jean-Luc Vila, "Optimal dynamic trading with leverage constraints," *Journal of Financial and Quantitative Analysis*, 1992, *27* (2), 151–168.

Guerrieri, Veronica and Guido Lorenzoni, "Credit crises, precautionary savings, and the liquidity trap," *The Quarterly Journal of Economics*, 2017, *132* (3), 1427–1467.

He, Zhiguo and Arvind Krishnamurthy, "A model of capital and crises," *The Review of Economic Studies*, 2012, *79* (2), 735–777.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante, "Consumption and labor supply with partial insurance: An analytical framework," *American Economic Review*, 2014, *104* (7), 2075–2126.

Herranz, Neus, Stefan Krasa, and Anne P Villamil, "Small firms in the SSBF," *Annals of Finance*, 2009, *5* (3), 341–359.

_ , _ , and _ , "Entrepreneurs, risk aversion, and dynamic firms," *Journal of Political Economy*, 2015, *123* (5), 1133–1176.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh, "The common factor in idiosyncratic volatility: Quantitative asset pricing implications," *Journal of Financial Economics*, 2016, *119* (2), 249–283.

_ , Thilo Kind, and Howard Kung, "Micro uncertainty and asset prices," *Available at SSRN 3220825*, 2018.

Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr, "Sources of US wealth inequality: Past, present, and future," *Nber macroeconomics annual*, 2021, *35* (1), 391–455.

Iachan, Felipe S, Dejanir Silva, and Chao Zi, "Under-diversification and idiosyncratic risk externalities," *Journal of Financial Economics*, 2021.

Jeong, Hyeok and Robert M Townsend, "Sources of TFP growth: occupational choice and financial deepening," *Economic Theory*, 2007, *32* (1), 179–221.

Kaplan, Greg, "Inequality and the life cycle," *Quantitative Economics*, 2012, *3* (3), 471–525.

_ and Giovanni L Violante, "How much consumption insurance beyond self-insurance?," *American Economic Journal: Macroeconomics*, 2010, *2* (4), 53–87.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry, "Agricultural decisions after relaxing credit and risk constraints," *The Quarterly Journal of Economics*, 2014, *129* (2), 597–652.

Kinnan, Cynthia and Robert Townsend, "Kinship and financial networks, formal financial access, and risk reduction," *American Economic Review*, 2012, *102* (3), 289–93.

Kiyotaki, Nobuhiro and John Moore, "Credit cycles," *Journal of Political economy*, 1997, *105* (2), 211–248.

Krueger, Dirk and Hanno Lustig, "When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)?," *Journal of Economic Theory*, 2010, *145* (1), 1–41.

Krusell, Per and Anthony A Smith Jr, "Income and wealth heterogeneity in the macroeconomy," *Journal of political Economy*, 1998, *106* (5), 867–896.

Kuznets, Simon, "Economic growth and income inequality," *The American economic review*, 1955, *45* (1), 1–28.

Ledoit, Olivier and Michael Wolf, "Honey, I shrunk the sample covariance matrix," *The Journal of Portfolio Management*, 2004, *30* (4), 110–119.

Lettau, Martin and Sydney Ludvigson, "Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying," *Journal of political economy*, 2001, *109* (6), 1238–1287.

Lucas, Robert E, "On the size distribution of business firms," *The Bell Journal of Economics*, 1978, pp. 508–523.

Mammen, Enno, "Bootstrap and wild bootstrap for high dimensional linear models," *The annals of statistics*, 1993, *21* (1), 255–285.

Mehra, Rajnish and Edward C. Prescott, "The equity premium: A puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Returns to capital in microenterprises: evidence from a field experiment," *The Quarterly Journal of Economics*, 2008, *123* (4), 1329–1372.

Merton, Robert C, "Lifetime portfolio selection under uncertainty: The continuous-time case," *The Review of Economics and Statistics*, 1969, pp. 247–257.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and misallocation: Evidence from plant-level data," *American Economic Review*, 2014, *104* (2), 422–58.

Moll, Benjamin, "Inequality and financial development: A power-law kuznets curve," *Princeton University Note*, 2012.

_ , "Productivity losses from financial frictions: Can self-financing undo capital misallocation?," *American Economic Review*, 2014, *104* (10), 3186–3221.

Morduch, Jonathan, "Income smoothing and consumption smoothing," *Journal of Economic Perspectives*, 1995, *9* (3), 103–114.

Moskowitz, Tobias J and Annette Vissing-Jørgensen, "The returns to entrepreneurial investment: A private equity premium puzzle?," *American Economic Review*, 2002, *92* (4), 745–778.

Onatski, Alexei, "Testing hypotheses about the number of factors in large factor models," *Econometrica*, 2009, *77* (5), 1447–1479.

Panousi, Vasia and Dimitris Papanikolaou, "Investment, idiosyncratic risk, and ownership," *The Journal of Finance*, 2012, *67* (3), 1113–1148.

Peter, Alessandra, "Equity Frictions and Firm Ownership," Working Paper 34301, National Bureau of Economic Research 2025.

Robinson, Baxter, "Risky business: The choice of entrepreneurial risk under incomplete markets," Technical Report, Working paper 2021.

Samphantharak, Krislert and Robert M Townsend, *Households as corporate firms: an analysis of household finance using integrated household surveys and corporate financial accounting* number 46, Cambridge University Press, 2010.

_ and _ , "Risk and return in village economies," *American Economic Journal: Microeconomics*, 2018, *10* (1), 1–40.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the Twenty-first Century," *The Quarterly Journal of Economics*, 2019, *134* (4), 1675–1745.

_ , Owen Zidar, Eric Zwick et al., "Top Wealth in America: New Estimates and Implications for Taxing the Rich," Technical Report 2020.

Stock, James H and Mark W Watson, "Macroeconomic forecasting using diffusion indexes," *Journal of Business & Economic Statistics*, 2002, *20* (2), 147–162.

Storesletten, Kjetil, Christopher I Telmer, and Amir Yaron, "Consumption and risk sharing over the life cycle," *Journal of monetary Economics*, 2004, *51* (3), 609–633.

_ , _ , and _ , "Asset pricing with idiosyncratic risk and overlapping generations," *Review of Economic Dynamics*, 2007, *10* (4), 519–548.

Sun, Yeneng, "The exact law of large numbers via Fubini extension and characterization of insurable risks," *Journal of Economic Theory*, 2006, *126* (1), 31–69.

Swanson, Eric T, "Risk aversion and the labor margin in dynamic equilibrium models," *American Economic Review*, 2012, *102* (4), 1663–91.

Tan, Eugene, "Entrepreneurial investment dynamics and the wealth distribution," 2018.

Townsend, Robert M, "Risk and insurance in village India," *Econometrica*, 1994, pp. 539–591.

Udry, Christopher and Santosh Anagol, "The return to capital in Ghana," *American Economic Review*, 2006, *96* (2), 388–393.

Viceira, Luis M, "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income," *The Journal of Finance*, 2001, *56* (2), 433–470.

Online Appendix

A Proofs

A.1 Proof of Lemma 1 and Proposition 1

Proof. We start by showing part (b) of Lemma 1, that is, we solve for $h_t(a)$ and its dynamics. Then, we proceed to solve for the entrepreneurs' value function and policy functions.

Pricing human wealth. Define the stochastic discount factor (SDF) for this economy as the process π_t satisfying the law of motion

$$\frac{d\pi_t}{\pi_t} = -r_t dt - p_t^{ag} dZ_t. \tag{A.1}$$

Without loss of generality, we assumed that the SDF is not exposed to idiosyncratic risk, as we only use the SDF to price human wealth which is not exposed to idiosyncratic risk. Integrating the process above, we obtain

$$\frac{\pi_z}{\pi_t} = \exp\left(-\int_t^z \left(r_u + \frac{(p_u^{ag})^2}{2}\right) du - \int_t^z p_u^{ag} dZ_u\right). \tag{A.2}$$

Similarly, integrating the process for A_t

$$\frac{A_z}{A_t} = \exp\left(\int_t^z \left(\mu_A - \frac{\sigma_A^2}{2}\right) du + \int_t^z \sigma_A dZ_u\right). \tag{A.3}$$

Hence, we can explicitly compute the following expectation

$$\begin{aligned}\mathbb{E}_t\left[\frac{\pi_z A_z}{\pi_t A_t}\right] &= \mathbb{E}_t\left[\exp\left(-\int_t^z \left(r_u - \mu_A + \frac{(p_u^{ag})^2 + \sigma_A^2}{2}\right) du - \int_t^z \left(p_u^{ag} - \sigma_A\right) dZ_u\right)\right] \\ &= \exp\left(-\int_t^z \left(r_u + p_u^{ag}\sigma_A - \mu_A\right) du\right),\end{aligned} \tag{A.4}$$

where we used Ito's isometry and the fact that p_t^{ag} is deterministic.

Human wealth is given by

$$h_t(a) = \mathbb{E}_t \left[\int_t^{t+T-a} \frac{\pi_z A_z}{\pi_t A_t} w_z \bar{l}(a+z-t) dz \right] = \int_t^{t+T-a} e^{-\int_t^z \left(r_u + p_u^{ag} \sigma_A - \mu_A \right) du} w_z \bar{l}(a+z-t) dz. \tag{A.5}$$

Consider the human wealth for someone born at date s, so $a = t - s$:

$$h_t(t-s) = \int_t^{s+T} e^{-\int_t^z \left(r_u + p_u^{ag} \sigma_A - \mu_A \right) du} w_z \bar{l}(z-s) dz. \tag{A.6}$$

Differentiating the expression above with respect to time yields

$$\frac{\partial h_t(a)}{\partial t} + \frac{\partial h_t(a)}{\partial a} = \left(r_t + p_t^{ag} \sigma_A - \mu_A \right) h_t(a) - w_t \bar{l}(a), \tag{A.7}$$

which gives (18) in a stationary equilibrium.

The HJB equation. The HJB equation for problem (13) is given by

$$\rho \tilde{V}_t(\tilde{n}, t-s; A_t) = \max_{\tilde{c}_t, \tilde{\theta}_t^{ag}, \tilde{\theta}_t^{id}, k_t, l_t, \iota_t} \left\{ \frac{\tilde{c}_t^{1-\gamma}}{1-\gamma} + \frac{\mathbb{E}_t \left[d\tilde{V}_{i,t} \right]}{dt} \right\}, \tag{A.8}$$

subject to (11) as well as the terminal and boundary conditions

$$\tilde{V}_t(\tilde{n}, T) = (1-\psi)^\gamma V^* \frac{\tilde{n}^{1-\gamma}}{1-\gamma}; \qquad \lim_{\tilde{n} \to -\tilde{h}_t(a)} \tilde{V}_t(\tilde{n}, a) = \begin{cases} 0, & \text{if } \gamma < 1 \\ -\infty, & \text{if } \gamma \geq 1 \end{cases}. \tag{A.9}$$

The terminal condition captures the effect of bequests and the boundary condition the fact that consumption is zero if the entrepreneur hits the natural borrowing limit.

Using Ito's lemma, the HJB reduces to a partial differential equation for $\tilde{V}_t(\tilde{n}, a; A_t)$:

$$\rho \tilde{V}_t = \max_{\tilde{c}_t, \tilde{\theta}_t^{ag}, \tilde{\theta}_t^{id}, k_t, l_t, \iota_t} \frac{\tilde{c}_t^{1-\gamma}}{1-\gamma} + \frac{\partial \tilde{V}_t}{\partial t} + \frac{\partial \tilde{V}_t}{\partial a} + \frac{\partial \tilde{V}_t}{\partial \tilde{n}} \mu_{\tilde{n},t} + \frac{\partial \tilde{V}_t}{\partial A_t} \mu_A A_t + \frac{1}{2} \frac{\partial^2 \tilde{V}_t}{\partial \tilde{n}^2} (\sigma_{ag,t}^2 + \sigma_{id,t}^2) + \frac{\partial^2 \tilde{V}_t}{\partial \tilde{n} \partial A_t} \sigma_{ag,t} \sigma_A A_t + \frac{1}{2} \frac{\partial \tilde{V}_i}{\partial A_t^2} \sigma_A^2 A_t^2, \tag{A.10}$$

subject to (11), where $(\mu_{\tilde{n},t}, \sigma_{ag,t}, \sigma_{id,t})$ are the drift and diffusion terms for $\tilde{n}_t$.

First, we verify that the following guess for the value function solves the PDE

$$\tilde{V}_t(\tilde{n}, a; A_t) = \zeta_t(a)^{-\gamma} \frac{\left(\tilde{n} + A_t h_t(a) \right)^{1-\gamma}}{1-\gamma}. \tag{A.11}$$

Plugging the derivatives of the equation above into the HJB equation, we obtain

$$\frac{\rho}{1-\gamma} = \max_{c_{i,t},k_{i,t},l_{i,t},\iota_{i,t},\theta^{ag}_{i,t},\theta^{id}_{i,t}} \Bigg\{ \frac{\zeta_t^{\gamma}(a)}{1-\gamma}\left(\frac{c_{i,t}}{\omega_{i,t}}\right)^{1-\gamma} - \frac{\gamma}{1-\gamma}\frac{1}{\zeta_t(a)}\left(\frac{\partial \zeta_t(a)}{\partial t} + \frac{\partial \zeta_t(a)}{\partial a}\right) + r_t + \frac{q_t k_{i,t}}{\omega_{i,t}}(\mu^R_{i,t} - r_t)$$
$$- \frac{p^{ag}_t \theta^{ag}_{i,t}}{\omega_{i,t}} + \frac{h_{i,t}}{\omega_{i,t}}\sigma_A p^{ag}_t - \frac{c_{i,t}}{\omega_{i,t}} - \frac{\gamma}{2}\left[\left(\frac{q_t k_{i,t} + h_{i,t}}{\omega_{i,t}}\sigma_A - \frac{\theta^{ag}_{i,t}}{\omega_{i,t}}\right)^2 + \left(\frac{q_t k_{i,t}}{\omega_{i,t}}\sigma_I - \frac{\theta^{id}_{i,t}}{\omega_{i,t}}\right)^2\right]$$
$$+ p^{id}_t \left[(1-\phi)\frac{q_t k_{i,t}}{\omega_{i,t}}\sigma_I - \frac{\theta^{id}_{i,t}}{\omega_{i,t}}\right]\Bigg\}, \tag{A.12}$$

where $p^{id}_{i,t}$ denotes the Lagrange multiplier on the skin-in-the-game constraint.

From the expression above, it is immediate that the optimal value of $(l_{i,t}, \iota_{i,t})$ maximizes the expected return on the business. The first-order conditions for $(l_{i,t}, \iota_{i,t})$ are given in (15) and (16), respectively. The expected return on the business will be equalized, allowing us to write $\mu^R_{i,t} = \mu^R_t$.

Policy functions. The first order condition for $\theta^{id}_{i,t}$ is given by

$$\gamma\left[\frac{q_t k_{i,t}}{\omega_{i,t}}\sigma_I - \frac{\theta^{id}_{i,t}}{\omega_{i,t}}\right] = p^{id}_{i,t}. \tag{A.13}$$

The equation above implies that the skin-in-the-game constraint is always binding, so $p^{id}_{i,t} > 0$ and $\theta^{id}_t = (1-\phi)q_t k_{i,t}\sigma_I$. If this was not the case, i.e. $p^{id}_{i,t} = 0$, then we would have $\theta^{id}_{i,t} = q_t k_{i,t}\sigma_I$, which violates the skin-in-the-game constraint.

The first-order conditions for capital and aggregate insurance are given by

$$\mu^R_t - r + p^{id}_{i,t}(1-\phi)\sigma_I = \gamma\left[\left(\frac{q_t k_{i,t} + h_{i,t}}{\omega_{i,t}}\sigma_A - \frac{\theta^{ag}_{i,t}}{\omega_{i,t}}\right)\sigma_A + \left(\frac{q_t k_{i,t}}{\omega_{i,t}}\sigma_I - \frac{\theta^{id}_{i,t}}{\omega_{i,t}}\right)\sigma_I\right]$$
$$p^{ag}_t = \gamma\left(\frac{q_t k_{i,t} + h_{i,t}}{\omega_{i,t}}\sigma_A - \frac{\theta^{ag}_{i,t}}{\omega_{i,t}}\right). \tag{A.14}$$

Combining the expressions above, we obtain

$$p^{id}_{i,t} = \frac{\mu^R_t - r_t - p^{ag}_t\sigma_A}{\phi\sigma_I}, \tag{A.15}$$

which coincides with expression (22) after we write $p^{id}_{i,t} = p^{id}_t$.

The demand for capital can be written as

$$\frac{q_t k_{i,t}}{\omega_{i,t}} = \frac{p_t^{id}}{\gamma\phi\sigma_I}. \tag{A.16}$$

Multiplying by $\omega_{i,t}/n_{i,t}$, we obtain expression (21). Solving for $\theta_{i,t}^{ag}$ in the optimality condition for aggregate insurance we obtain (23).

The first-order condition for consumption gives

$$\frac{c_{i,t}}{\omega_{i,t}} = \zeta_t(a). \tag{A.17}$$

Plugging the expressions above back into the HJB, we obtain a PDE for $\zeta_t(a)$

$$\frac{\partial\zeta_t(a)}{\partial t} + \frac{\partial\zeta_t(a)}{\partial a} = \zeta_t^2(a) - \bar{r}_t\zeta_t(a), \tag{A.18}$$

where $\bar{r}_t \equiv \frac{1}{\gamma}\rho + \left(1 - \frac{1}{\gamma}\right)\left[r + \frac{(p_t^{id})^2 + (p_t^{ag})^2}{2\gamma}\right]$.

Define $z_{s,t} \equiv \zeta_t^{-1}(t-s)$ as the wealth-consumption ratio for an entrepreneur born at date s. Differentiating with respect to t, we obtain

$$\dot{z}_{s,t} = -\frac{1}{\zeta_t^2(a)}\left[\frac{\partial\zeta_t}{\partial t} + \frac{\partial\zeta_t}{\partial a}\right] = \bar{r}_t z_{s,t} - 1. \tag{A.19}$$

Solving the above differential equation, we get

$$z_{s,t} = \int_t^{s+T} e^{-\int_t^u \bar{r}_z dz} du + e^{-\int_t^{s+T} \bar{r}_z dz} z_{s,s+T}, \tag{A.20}$$

or in terms of $\zeta_t(a)$, we have

$$\zeta_t(a) = \frac{1}{\int_t^{t+T-a} e^{-\int_t^u \bar{r}_z dz} du + e^{-\int_t^{t+T-a} \bar{r}_z dz}(1-\psi)(V^*)^{\frac{1}{\gamma}}}, \tag{A.21}$$

where we used the boundary condition $\zeta_t^{-1}(T) = (1-\psi)\,(V^*)^{\frac{1}{\gamma}}$.

Assuming $(V^*)^{\frac{1}{\gamma}} = \frac{1}{\bar{r}}$ and a stationary equilibrium, where $\bar{r}_t = \bar{r}$, we obtain

$$\zeta(a) = \frac{\bar{r}}{1 - \psi e^{-\bar{r}(T-a)}}. \tag{A.22}$$

$(V^*)^{\frac{1}{\gamma}} = \frac{1}{\bar{r}}$ ensures $\psi = 0$ gives the corresponding value in an infinite-horizon economy.

Price and quantity of aggregate insurance. The demand for aggregate insurance for entrepreneus is $\theta^{ag}_{i,t} = (q_t k_{i,t} + h_{i,t})\sigma_A - \omega_{i,t}\frac{p^{ag}_t}{\gamma}$. A similar expression holds for wage earners

$$\theta^{ag}_{j,t} = h_{j,t}\sigma_A - (n_{j,t} + h_{j,t})\frac{p^{ag}}{\gamma}, \tag{A.23}$$

Combining the demand for aggregate insurance for entrepreneurs and wage earners with the corresponding market-clearing condition, we obtain

$$\int_{\mathcal{E}_t}\left[(qk_{i,t} + h_{i,t})\sigma_A - (n_{i,t} + h_{i,t})\frac{p^{ag}}{\gamma}\right]di + \int_{\mathcal{W}_t}\left[h_{j,t}\sigma_A - (n_{j,t} + h_{j,t})\frac{p^{ag}}{\gamma}\right]di = 0. \tag{A.24}$$

Rearranging the expression above, we can solve for the price of aggregate insurance p^{ag}

$$p^{ag}_t = \frac{\int_{\mathcal{E}_t}(qk_{i,t} + h_{i,t})di + \int_{\mathcal{W}_t} h_{j,t}dj}{\int_{\mathcal{E}_t}(n_{i,t} + h_{i,t})di + \int_{\mathcal{W}_t}(n_{j,t} + h_{j,t})dj}\gamma\sigma_A = \gamma\sigma_A, \tag{A.25}$$

using the fact that $\int_{\mathcal{E}_t} n_{i,t}di + \int_{\mathcal{W}_t} n_{j,t}dj = \int_{\mathcal{E}_t} qk_{i,t}di$. The demand for aggregate insurance for entrepreneurs if then given by $\theta^{ag}_{j,t} = (q_t k_{i,t} - n_{i,t})\sigma_A$.

□

A.2 Proof of Proposition 2

Proof. We start by deriving the law of motion of financial wealth for an entrepreneur of a given age. Using the value of capital, $k_{i,t}$, aggregate and idiosyncratic insurance, $(\theta^{ag}_{i,t}, \theta^{id}_{i,t})$, and the definition of the price of idiosyncratic risk, p^{id}_t, given in Proposition 1, we can write the law of motion of financial wealth as follows

$$\begin{aligned} d\tilde{n}_{i,t} = & \left[r_t\tilde{\omega}_{i,t} + \frac{(p^{id}_t)^2}{\gamma}\tilde{\omega}_{i,t} + \frac{(p^{ag}_t)^2}{\gamma}\tilde{\omega}_{i,t} - \tilde{h}_{i,t}(r_t + p^{ag}_t\sigma_A) + \tilde{w}_t\bar{l}_{i,t} - \tilde{c}_{i,t}\right]dt \\ & + \left(\tilde{\omega}_{i,t}\frac{p^{ag}_t}{\gamma} - \tilde{h}_{i,t}\sigma_A\right)dZ_t + \frac{p^{id}_t}{\gamma}\tilde{\omega}_{i,t}dZ_{i,t}, \end{aligned} \tag{A.26}$$

where $\tilde{\omega}_{i,t} = \tilde{n}_{i,t} + \tilde{h}_{i,t}$.

Using the fact that $p^{ag}_t = \gamma\sigma_A$ in equilibrium, we find that the aggregate risk exposure of entrepreneurs is given $\tilde{n}_{i,t}\sigma_A$. Hence, scaled financial wealth, $n_{i,t} = \tilde{n}_{i,t}/A_t$, does not respond to aggregate shocks. The evolution of $n_{i,t}$ can then be written as

$$dn_{i,t} = \mu_{n,t}(n_{i,t}, a)dt + \sigma_{n,t}(n_{i,t}, a)dZ_{i,t}, \tag{A.27}$$

where

$$\mu_{n,t}(n,a) = \left[r_t + \frac{(p_t^{id})^2}{\gamma} + \frac{(p_t^{ag})^2}{\gamma} - \mu_A - \zeta_t(a)\right](n + h_t(a)) - \mu_{h,t}(a) \tag{A.28}$$

$$\sigma_{n,t}(n,a) = \frac{p_t^{id}}{\gamma}(n + h_t(a)), \tag{A.29}$$

and $\mu_{h,t}(a)$ is the drift of $h_t(a)$.

Derivation of Equation (25). Notice that total wealth evolves according to

$$\frac{d\omega_{i,t}}{\omega_{i,t}} = \left[r_t + \frac{(p_t^{ag})^2}{\gamma} + \frac{(p_t^{id})^2}{\gamma} - \mu_A - \zeta_t(t - s_i)\right]dt + \frac{p_t^{id}}{\gamma}dZ_{i,t}, \tag{A.30}$$

where s_i is the birthdate of entrepreneur i.

Let $\overline{\omega}_{s,t} \equiv \frac{\int_{\mathcal{E}_t} \mathbb{1}_{s_i=s}\omega_{i,t}di}{\int_{\mathcal{E}_t} \mathbb{1}_{s_i=s}di}$ denote the average total wealth of entrepreneurs born at date s. The law of motion of $\overline{\omega}_{s,t}$ is given by

$$d\overline{\omega}_{s,t} = \left[r_t + \frac{(p_t^{ag})^2}{\gamma} + \frac{(p_t^{id})^2}{\gamma} - \mu_A - \zeta_t(t - s)\right]\overline{\omega}_{s,t}dt \tag{A.31}$$

where the idiosyncratic risk is diversified by averaging out across entrepreneurs of a given cohort.

It is convenient to express total wealth as a function of age instead of the entrepreneurs' birthdate. Let $\omega_t(a)$ denote the average total wealth of investors with age a at period t. Using the fact that $\overline{\omega}_{s,t} = \omega_t(t-s)$, we obtain the following PDE for $\omega_t(a)$:

$$\frac{\partial\omega_t(a)}{\partial t} + \frac{\partial\omega_t(a)}{\partial a} = \left[r_t + \frac{(p_t^{ag})^2}{\gamma} + \frac{(p_t^{id})^2}{\gamma} - \mu_A - \zeta_t(a)\right]\omega_t(a). \tag{A.32}$$

In a stationary equilibrium, $\omega_t(a)$ does not depend on calendar time t, which allow us to write

$$\frac{d\log\omega(a)}{da} = r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \mu_A - \zeta(a). \tag{A.33}$$

Integrating the expression above, we obtain

$$\log\omega(a) = \log\omega(0) + \left[r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \mu_A\right]a - \int_0^a \zeta(u)du. \tag{A.34}$$

Using the fact $\log\frac{\omega(a)}{\omega(0)} = \log\frac{f(a)\omega(a)}{f(0)\omega(0)} + ga$ and the identity $\omega(a) = n(a)\left(1+\frac{h(a)}{n(a)}\right)$, we obtain expression (25) after some rearrangement.

Initial wealth. The expression for $\omega(a)$ in levels can be written as

$$\omega(a) = \omega(0)e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-\mu_A\right)a}\frac{e^{-\bar{r}a}-\psi e^{-\bar{r}T}}{1-\psi e^{-\bar{r}T}}. \tag{A.35}$$

Evaluating at $a = T$ gives

$$\omega(T) = \omega(0)e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-\mu_A-mpc_e\right)T}, \tag{A.36}$$

where $mpc_e = \frac{1}{T}\int_0^T \zeta(a)da$.

The boundary condition at age T implies $\omega(0) = e^{-gT}\omega(T) + h(0)$, then

$$\omega(0) = \frac{h(0)}{1-e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-\mu_A-g-mpc_e\right)T}}. \tag{A.37}$$

Using $\omega(0) = n(0) + h(0)$ and rearranging the resulting expression for $n(0)$.

Derivation of Equation (26). Multiplying Equation (A.35) by $f(a)$, integrating over age, and using the fact that $f(a) = e^{-ga}f(0)$, we obtain

$$n_e + h_e = f(0)\omega(0)\int_0^T e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-(g+\mu_A)\right)a}\frac{e^{-\bar{r}a}-\psi e^{-\bar{r}T}}{1-\psi e^{-\bar{r}T}}da, \tag{A.38}$$

which gives Equation (26) after some rearrangement .

□

A.3 Proof of Lemma 2

Proof. We derive the Kolmogorov Forward Equation as the limit of a discrete-time economy. The discrete-time approximation goes as follows. Time takes values on the discrete set $\{t^1,\ldots,t^L\}$, where $\Delta t = t^{l+1}-t^l$ is the constant time step. Scaled financial wealth $n_{i,t}$ takes values on a discrete grid, $n_{i,t}\in\{n^1,n^2,\ldots,n^J\}$ with a constant step size $\Delta n = n^{j+1}-n^j$. Age is also assumed to take values in a discrete grid $\{a^1,\ldots,a^K\}$, where $\Delta a = a^{k+1}-a^k$, $a^1 = 0$, and $a^K = T$. For simplicity, assume $\Delta a = \Delta t$. The probability of moving up, down, or staying at the same point of the grid is chosen to approximate (A.27) and are given, respectively, by

$$p_u(n^j, a^k) = \frac{1}{2}\left[\frac{\sigma_n(n^j, a^k)^2}{\overline{\sigma}^2} + \frac{\mu_n(n^j, a^k)}{\overline{\sigma}^2}\Delta n\right] \tag{A.39}$$

$$p_d(n^j, a^k) = \frac{1}{2}\left[\frac{\sigma_n(n^j, a^k)^2}{\overline{\sigma}^2} - \frac{\mu_n(n^j, a^k)}{\overline{\sigma}^2}\Delta n\right] \tag{A.40}$$

$$p_s(n^j, a^k) = 1 - \frac{\sigma_n(n^j, a^k)^2}{\overline{\sigma}^2}. \tag{A.41}$$

where $\overline{\sigma} = \max_{1\leq j\leq J, 1\leq k\leq K} \sigma_n(n^j, a^k)$, $\Delta n = \overline{\sigma}\sqrt{\Delta t}$, and $\Delta a = \Delta t$.

Notice that the expected change in $n_{i,t}$, where $n_{i,t} = n^j$ and $a_i = a^k$, is given by

$$\mathbb{E}\left[n_{i,t+1} - n_{i,t}\right] = p_u(n^j, a^k)\Delta n + p_d(n^j, a^k)(-\Delta n) = \mu_n(n^j, a^k)\Delta t, \tag{A.42}$$

and

$$\mathbb{E}\left[(n_{i,t+1} - n_{i,t})^2\right] = p_u(n^j, a^k)\Delta n^2 + p_d(n^j, a^k)(-\Delta n)^2 = \sigma_n(n^j, a^k)^2\Delta t. \tag{A.43}$$

Let $m(n^j, a^k, t^l)$ denote the mass of agents with financial wealth n^j, age a^k, at period t^l. Summing over n^j, we obtain the mass of agents with age a^k, $M_{k,l} \equiv \sum_{j=1}^{J} m(n^j, a^k, t^l) = e^{g(t^l-(k-1)\Delta t)}$. Summing over (n^j, a^k), we obtain the total population $M_l = \sum_{k=1}^{K} M_{k,l}$, so $M_{l+1} = e^{g\Delta t}M_l$. The law of motion of m, for $k > 1$ and $1 < j < J$, is given by

$$\begin{aligned} m(n^j, a^k, t^l + \Delta t) = {} & p_u(n^j - \Delta n, a^k - \Delta a)m(n^j - \Delta n, a^k - \Delta a, t^l) + p_s(n^j, a^k - \Delta a)m(n^j, a^k - \Delta a, t^l) \\ & + p_d(n^j + \Delta n, a^k - \Delta a)m(n^j + \Delta n, a^k - \Delta a, t^l). \end{aligned} \tag{A.44}$$

The boundary conditions are defined as follows. For $j = 1$ and $j = J$, we will assume a reflecting boundary, that is, if n moves up from n_J or down from n_1, it is immediately reflected back to its initial position

$$\begin{aligned} m(n^J, a^k, t^l + \Delta t) = {} & p_u(n^J - \Delta n, a^k - \Delta a)m(n^J - \Delta n, a^k - \Delta a, t^l) + p_s(n^J, a^k - \Delta a)m(n^J, a^k - \Delta a, t^l) \\ & + p_u(n^j, a^k - \Delta a)m(n^J, a^k - \Delta a, t^l), \end{aligned} \tag{A.45}$$

and analogously for $j = 1$.

Finally, for $k = 1$, we have

$$m(e^{-gT}n^j, a^1, t^l + \Delta t) = e^{gT}\left[p_u\left(n^j - \Delta n, a^K\right) m\left(n^j - \Delta n, a^K, t^l\right) + p_s\left(n^j, a^k\right) m\left(n^j, a^K, t^l\right) + p_d\left(n^j + \Delta n, a^K\right) m\left(n^j + \Delta n, a^K, t^l\right)\right]. \tag{A.46}$$

since each one of the e^{gT} heirs inherit $e^{-gT}n^j$, where we assumed $e^{-gT}n^j$ belongs to the grid.

Let $f(n^j, a^k, t^l) \equiv \frac{m(n^j, a^k, t^l)}{M_l}$ denote the share of agents in state (n^j, a^k) in period t^l. Dividing both sides of (A.44) by M_l and taking a Taylor expansion, we obtain

$$\begin{aligned}(1+g\Delta t)(f + f_t\Delta t) = &\frac{1}{2}\left(\frac{\sigma_n^2 - (\sigma_n^2)_n\Delta n + 0.5(\sigma_n^2)_{nn}\Delta n^2 - (\sigma_n^2)_a\Delta t}{\overline{\sigma}^2} + \frac{\mu_n - (\mu_n)_n\Delta n}{\overline{\sigma}^2}\Delta n\right)\left(f - f_a\Delta t - f_n\Delta n + 0.5 f_{nn}\Delta n^2\right) \\ &+ \frac{1}{2}\left(\frac{\sigma_n^2 + (\sigma_n^2)_n\Delta n + 0.5(\sigma_n^2)_{nn}\Delta n^2 - (\sigma_n^2)_a\Delta t}{\overline{\sigma}^2} - \frac{\mu_n + (\mu_n)_n\Delta n}{\overline{\sigma}^2}\Delta n\right)\left(f - f_a\Delta t + f_n\Delta n + 0.5 f_{nn}\Delta n^2\right) \\ &+ \left(1 - \frac{\sigma_n^2 - (\sigma_n^2)_a\Delta t}{\overline{\sigma}^2}\right)(f - f_a\Delta t) + o(\Delta t).\end{aligned} \tag{A.47}$$

Simplifying the expression above and taking the limit $\Delta t \to 0$, we obtain

$$f_t + f_a + gf = \frac{1}{2}(\sigma_n^2)_{nn} f - (\mu_n)_n f + (\sigma_n^2)_n f_n - \mu_n f_n + \frac{1}{2}\sigma_n^2 f_{nn}, \tag{A.48}$$

or, more explicitly, we can write the expression as follows

$$\frac{\partial f(n,a,t)}{\partial t} + \frac{\partial f(n,a,t)}{\partial a} + gf(n,a,t) = -\frac{\partial[f(n,a,t)\mu_n(n,a)]}{\partial n} + \frac{1}{2}\frac{\partial\left[f(n,a,t)\sigma_n^2(n,a)\right]}{\partial n^2}. \tag{A.49}$$

Let $f_t(n|a)$ denote the conditional density at date t, so $f_t(n,a) = f_t(n|a)f(a)$. We can write the Kolmogorov Forward Equation in terms of the conditional density:

$$f(a)\frac{\partial f_t(n|a)}{\partial t} + f(a)\frac{\partial f_t(n|a)}{\partial a} + f_t(n|a)f'(a) = -f(a)\frac{\partial[f_t(n|a)\mu_n(n,a)]}{\partial n} + f(a)\frac{1}{2}\frac{\partial\left[f_t(n|a)\sigma_n^2(n,a)\right]}{\partial n^2} - gf_t(n,a). \tag{A.50}$$

Dividing by $f(a)$ and using the fact that $f'(a) = -gf(a)$, we obtain

$$\frac{\partial f_t(n|a)}{\partial t} + \frac{\partial f_t(n|a)}{\partial a} = -\frac{\partial[f_t(n|a)\mu_n(n,a)]}{\partial n} + \frac{1}{2}\frac{\partial\left[f_t(n|a)\sigma_n^2(n,a)\right]}{\partial n^2}. \tag{A.51}$$

In a stationary equilibrium, we can ignore the dependence on calendar time to obtain

$$\frac{\partial f(n|a)}{\partial a} = -\frac{\partial[f(n|a)\mu_n(n,a)]}{\partial n} + \frac{1}{2}\frac{\partial\left[f(n|a)\sigma_n^2(n,a)\right]}{\partial n^2}. \tag{A.52}$$

□

A.4 Proof of Proposition 3

Proof. The law of motion of (log) total wealth is

$$d\log\omega_{i,t} = \left[r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \frac{\bar{r}}{1-e^{-\bar{r}(T-(t-s))}} - \mu_A - \frac{1}{2}\left(\frac{p^{id}}{\gamma}\right)^2\right]dt + \frac{p^{id}}{\gamma}dZ_{i,t}, \tag{A.53}$$

where s denotes the birth date of entrepreneur i.

Integrating the expression above, we obtain

$$\log\omega_{i,t} = \log\omega_{i,s} + \int_s^t \left[r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \frac{\bar{r}}{1-e^{-\bar{r}(T-(t'-s))}} - \mu_A - \frac{1}{2}\left(\frac{p^{id}}{\gamma}\right)^2\right]dt' + \frac{p^{id}}{\gamma}(Z_{i,t}-Z_{i,s}), \tag{A.54}$$

where $Z_{i,t} - Z_{i,s} \sim \mathcal{N}(0,a)$ and $a = t - s$.

Hence, $\log\omega_{i,t} \sim \mathcal{N}(m(a), v(a))$, where the mean and variance are given by

$$m(a) = \log h(0) + \left[r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \mu_A - \frac{1}{2}\left(\frac{p^{id}}{\gamma}\right)^2 - \bar{r}\right]a + \log\frac{1-e^{-\bar{r}(T-a)}}{1-e^{-\bar{r}T}} \tag{A.55}$$

$$v(a) = \left(\frac{p^{id}}{\gamma}\right)^2 a, \tag{A.56}$$

using the fact that $\omega_{i,s_i} = h(0)$ when $\psi = 1$.

As the ratio of consumption to total wealth is the same for all entrepreneurs of the same age, the variance of log consumption is given by

$$\mathbb{V}[\log c_{i,t}|a] = \mathbb{V}[\log\omega_{i,t}|a] = \left(\frac{p^{id}}{\gamma}\right)^2 a. \tag{A.57}$$

Normalized financial wealth $n_{i,t} = \omega_{i,t} - h_{i,t}$ has a shifted log-normal distribution conditional on $s_i = s$, with support $(-h(a), \infty)$. The expected value and variance of $n_{i,t}$ are

$$\mathbb{E}[n|a] = h(0)e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-\mu_A-\bar{r}\right)a}\frac{1-e^{-\bar{r}(T-a)}}{1-e^{-\bar{r}T}} - h(a) \tag{A.58}$$

$$\mathbb{V}[n|a] = \left[e^{\left(\frac{p^{id}}{\gamma}\right)^2 a} - 1\right]\left[h(0)e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-\mu_A\right)a}\frac{e^{-\bar{r}a}-e^{-\bar{r}T}}{1-e^{-\bar{r}T}}\right]^2. \tag{A.59}$$

We show next that $\mathbb{V}[n|a]$ has an inverted U shape. Define the following functions:

$$v_1(a) = \left[e^{\left(\frac{p^{id}}{\gamma}\right)^2 a} - 1\right]^{\frac{1}{2}} e^{\left(r+\frac{(p^{ag})^2}{\gamma}+\frac{(p^{id})^2}{\gamma}-\mu_A\right)a}; \qquad v_2(a) = \frac{e^{-\bar{r}a}-e^{-\bar{r}T}}{1-e^{-\bar{r}T}}. \tag{A.60}$$

The derivative of the product of $v_1(a)$ and $v_2(a)$ will be positive if

$$v_1'(a)v_2(a) + v_1(a)v_2'(a) > 0 \iff \frac{v_1'(a)}{v_1(a)} > -\frac{v_2'(a)}{v_2(a)}, \tag{A.61}$$

for $a \neq 0$ and $a \neq T$. Notice that $-v_2'(a)/v_2(a)$ is positive, monotonically increasing, and approaches ∞ as a approaches T:

$$-\frac{v_2'(a)}{v_2(a)} = \bar{r}\frac{1}{1 - e^{-\bar{r}(T-a)}}. \tag{A.62}$$

The term $v_1'(a)/v_1(a)$ is positive, monotonically decreasing, and approaches $+\infty$ as $a \to 0$:

$$\frac{v_1'(a)}{v_1(a)} = \frac{1}{2}\left(\frac{p^{id}}{\gamma}\right)^2 \frac{e^{\left(\frac{p^{id}}{\gamma}\right)^2 a}}{e^{\left(\frac{p^{id}}{\gamma}\right)^2 a} - 1} + r + \frac{(p^{ag})^2}{\gamma} + \frac{(p^{id})^2}{\gamma} - \mu_A. \tag{A.63}$$

Hence, there exists a unique $0 < \hat{a} < T$ such that $v_1'(a)v_2(a) + v_1(a)v_2'(a) > 0$ for all $a < \hat{a}$ and $v_1'(a)v_2(a) + v_1(a)v_2'(a) < 0$ for all $a > \hat{a}$. Hence, $\mathbb{V}[n|a]$ is equal to zero at $a = 0$, it increases monotonically for $a < \hat{a}$, where it achieves the maximum, and it decreases towards zero for $\hat{a} < a \leq T$.

□

A.5 Proof of Lemma 3

Proof. The HJB for the entrepreneur's problem is given by

$$\rho \tilde{V}_t = \max_{\tilde{c}_i, \tilde{\theta}_i^{ag}, \tilde{\theta}_t^{id}, k_i, l_i, \iota_i} \frac{\tilde{c}_t^{1-\gamma}}{1-\gamma} + \frac{\mathbb{E}_t\left[d\tilde{V}_t\right]}{dt} \tag{A.64}$$

subject to (11). We guess-and-verify that the value function can be written as

$$\tilde{V}^j(\tilde{\omega}; A_t) = A_t^{1-\gamma} V^j\left(\frac{\tilde{\omega}}{A_t}\right). \tag{A.65}$$

where $V_j(\omega)$ is independent of A_t, and $j \in \{1, \ldots, n\}$ index the entrepreneur's type.

Let $\omega_{i,t} \equiv \tilde{\omega}_{i,t}/A_t$ denote scaled total wealth. From the law of motion of $h_{i,t} = \tilde{h}_{i,t}/A_t$ in Lemma 1 and the law of motion of $n_{i,t}$, we obtain

$$d\omega_{i,t} = \mu_{\omega_i,t}dt + \sigma_{i,t}^{ag}dZ_t + \sigma_{i,t}^{id}dZ_{i,t}, \tag{A.66}$$

where $\sigma^{ag}_{i,t} \equiv (qk_{i,t} + h_{i,t} - \omega_{i,t})\,\sigma_A - \theta^{ag}_{i,t}$, $\sigma^{id}_{i,t} \equiv qk_{i,t}\sigma_{I,i} - \theta^{id}_{i,t}$, and

$$\mu_{\omega_i,t} \equiv (r + p^{ag}\sigma_A - \mu_A)\omega_{i,t} + qk_{i,t}(\hat{\mu}^R_{i,t} - r - p^{ag}\sigma_A) + (p^{ag} - \sigma_A)\sigma^{ag}_{i,t} - c_{i,t}. \tag{A.67}$$

The HJB equation for the scaled value function can be written as

$$\hat{\rho}V^j = \max_{c_i, \theta^{ag}_i, \theta^{id}_t, k_i, l_i, \iota_i} \frac{c_t^{1-\gamma}}{1-\gamma} + V^j_a + V^j_\omega \left[\hat{r}\omega_{i,t} + qk_{i,t}(\hat{\mu}^R_{i,t} - r - p^{ag}\sigma_A) + \hat{p}^{ag}\sigma^{ag}_{i,t} - c_{i,t}\right] + \frac{1}{2}V^j_{\omega\omega}\left((\sigma^{ag}_{i,t})^2 + (\sigma^{id}_{i,t})^2\right). \tag{A.68}$$

subject to $\theta^{id}_{i,t} \le (1-\phi)qk_{i,t}\sigma_{I,i}$ and $\omega_{i,t} \ge 0$, where $\sigma_{I,i} = \sigma^j_I$ and

$$\hat{\rho} \equiv \rho - (1-\gamma)\left(\mu_A - \frac{\gamma\sigma_A^2}{2}\right), \qquad \hat{r} \equiv r + p^{ag}\sigma_A - \mu_A, \qquad \hat{p}^{ag} \equiv p^{ag} - \gamma\sigma_A. \tag{A.69}$$

It is optimal to choose $l_{i,t}$ and $\iota_{i,t}$ to maximize expected returns. The investment demand is then given by Equation (16) and the labor demand is given by $w = \beta k^{\alpha}_{i,t} l^{\beta-1}_{i,t}$.

The expected return for entrepreneur i is given by

$$\hat{\mu}^R_{i,t} = \frac{(1-\beta)\left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} k_{i,t}^{\frac{\alpha+\beta-1}{1-\beta}} - \iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta. \tag{A.70}$$

The first-order condition for aggregate insurance is given by

$$\hat{p}^{ag}_t = -\frac{V^j_{\omega\omega}}{V^j_\omega}\sigma^{ag}_{i,t} \Rightarrow \theta^{ag}_{i,t} = (qk_{i,t} - n_{i,t})\sigma_A + \frac{V^j_\omega}{V^j_{\omega\omega}}\hat{p}^{ag}. \tag{A.71}$$

In equilibrium, we must have that average $\sigma^{ag}_{i,t}$ across entrepreneurs and workers must be equal to zero, such that scaled wealth is, on average, not exposed to aggregate risk. Then, we have $\hat{p}^{ag}_t = 0$, so $p^{ag}_t = \gamma\sigma_A$, as in the baseline model. The first-order conditions for $k_{i,t}$ and $\theta^{id}_{i,t}$ are given by

$$\mu^R_{i,t} - r - \gamma\sigma_A^2 + p^{id}_i(1-\phi)\sigma_{I,i} = -\frac{V^j_{\omega\omega}}{V^j_\omega}\left[\sigma^{ag}_{i,t}\sigma_A + \sigma^{id}_{i,t}\sigma_{I,i}\right], \qquad p^{id}_{i,t} = -\frac{V^j_{\omega\omega}}{V^j_\omega}\sigma^{id}_{i,t}, \tag{A.72}$$

where $\mu^R_{i,t}$ denotes the expected *marginal* return on the business:

$$\mu^R_{i,t} = \frac{\alpha\left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} k_{i,t}^{\frac{\alpha+\beta-1}{1-\beta}} - \iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta. \tag{A.73}$$

$\mu^R_{i,t}$ corresponds to the expected profit from an additional unit of capital divided by the cost of one unit of capital, while $\hat{\mu}^R_{i,t}$ corresponds to the expected total profit divided by the cost of the entire capital stock. Under constant returns to scale, the two concepts coincide. With DRS, the marginal concept is the relevant determining the investment decision.

Given that $\sigma^{id}_{i,t} > 0$ by the insurance constraint, and given the concavity of the value function, $V^j_{\omega\omega} < 0$, we have that $p^{id}_{i,t} > 0$. Therefore, the insurance constraint is always binding, that is, $\theta_{i,t} = (1-\phi)qk_{i,t}\sigma_{I,i}$. Rearranging the expressions above, we obtain

$$p^{id}_{i,t} = \frac{\mu^R_{i,t} - r - \gamma\sigma^2_A}{\phi\sigma_{I,i}}, \qquad qk_{i,t} = -\frac{V^j_\omega}{V^j_{\omega\omega}}\frac{p^{id}_{i,t}}{\phi\sigma_{I,i}}. \tag{A.74}$$

Finally, the optimality condition for consumption is given by $c^{-\gamma}_{i,t} = V^j_\omega$. □

A.6 Proof of Propositions 4 and 5

Proof. We consider the case where entrepreneurs have different levels of skill, $e_i \in \{e^1, \ldots, e^n\}$, and different exposures to idiosyncratic risk $\sigma_{I,i} \in \{\sigma^1_I, \ldots, \sigma^n_I\}$. The share of entrepreneurs of type j is denoted by $\chi_{e,j}$. Consider the log-linear approximation of the policy functions:

$$\hat{c}_{i,t} = \psi^j_{c,\omega}\hat{\omega}_{i,t}, \qquad \hat{k}_{i,t} = \psi^j_{k,\omega}\hat{\omega}_{i,t}, \tag{A.75}$$

up to first order in $\hat{w}_{i,t} = \log\omega_{i,t} - \log\overline{\omega}_j$, where $\overline{\omega}_j \equiv \exp\mathbb{E}[\log\omega_{i,t}|\sigma^i_I = \sigma^j_I, e_i = e^j]$, $\hat{c}_{i,t} = \log c_{i,t} - \log c_j(\overline{\omega}_j)$, and $\hat{k}_{i,t} = \log k_{i,t} - \log k_j(\overline{\omega}_j)$.

Wages. Aggregating the labor demand condition, $w_t = \beta k^\alpha_{i,t} l^{\beta-1}_{i,t} e^{1-\alpha-\beta}_i$, we obtain

$$\int_{\mathcal{E}} l_{i,t}di = \left(\frac{\beta}{w_t}\right)^{\frac{1}{1-\beta}} \int_{\mathcal{E}} k^{\frac{\alpha}{1-\beta}}_{i,t} e^{\frac{1-\alpha-\beta}{1-\beta}}_{i,t} di = \left(\frac{\beta}{w_t}\right)^{\frac{1}{1-\beta}} \sum_{j=1}^n \tilde{\chi}_{e,j} e_i^{\frac{1-\alpha-\beta}{1-\beta}} \bar{k}_j^{\frac{\alpha}{1-\beta}} \left[\chi_e + \frac{\alpha\psi_{k,\omega}}{1-\beta}\frac{1}{\tilde{\chi}_{e,j}}\int_{\mathcal{E}_j}\hat{\omega}_{i,t}di\right]. \tag{A.76}$$

Using a law of large numbers to obtain $\int\hat{\omega}_{i,t}di = 0$, we obtain

$$w_t = \beta K^\alpha \left(\frac{\chi_e}{\bar{l}}\right)^{1-\alpha-\beta}, \tag{A.77}$$

where $K = \chi_e\left[\sum_{j=1}^n \tilde{\chi}_{e,j}(\bar{k}^j)^{\frac{\alpha}{1-\beta}}\right]^{\frac{1-\beta}{\alpha}}/\bar{l}$ is the aggregate capital-labor ratio, and $\tilde{\chi}_{e,j} \equiv \chi_{e,j}e_i^{\frac{1-\alpha-\beta}{1-\beta}}$. We choose the units of skill such that $\sum_{j=1}^n \tilde{\chi}_{e,j} = 1$. We also adopt the normalization $\bar{l} = \chi_e$, so $w_t = \beta K^\alpha$.

Expected returns. Expected returns are given by

$$\hat{\mu}^R_{i,t} = \overline{\hat{\mu}}^R + \psi_{\hat{\mu}^R,k}\hat{k}_{i,t}, \tag{A.78}$$

where $\overline{\hat{\mu}}^R \equiv \frac{(1-\beta)B\bar{k}^{\alpha-1}-\iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta$ and $\psi_{\hat{\mu}^R,k} \equiv -\frac{(1-\alpha-\beta)B\bar{k}^{\alpha-1}}{q}$, where $B_j = e_j^{\frac{1-\alpha-\beta}{1-\beta}}\left(\frac{\bar{k}_j}{K}\right)^{\frac{\alpha\beta}{1-\beta}}$. To ease notation, we dropped the subscript j from above.

Expected marginal returns. The expected marginal return on the business is given by

$$\mu^R_{i,t} = \overline{\mu}^R - \psi_{\mu^R,k}\hat{k}_t, \tag{A.79}$$

where $\overline{\mu}^R \equiv \frac{\alpha B\bar{k}^{\alpha-1}-\iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta$ and $\psi_{\mu^R,k} \equiv \frac{1-\alpha-\beta}{1-\alpha}\frac{\alpha B\bar{k}^{\alpha-1}}{q}$.

Log portfolio returns. Define the expected log portfolio returns as

$$r^p_{i,t} \equiv \hat{r} + z_{i,t}(\mu^R_{i,t} - r - \gamma\sigma_A^2) - \frac{1}{2}\left(z_{i,t}\phi\sigma_I\right)^2. \tag{A.80}$$

where $z_{i,t} \equiv \frac{qk_{i,t}}{\omega_{i,t}}$. Log-linearizing the expression above, we obtain

$$r^p_{i,t} = \bar{r}^p + \psi_{r^p,z}\hat{z}_{i,t} + \psi_{r^p,k}k_{i,t}, \tag{A.81}$$

$\bar{r}^p \equiv \hat{r} + \bar{z}(\overline{\hat{\mu}}^R - r - \gamma\sigma_A^2) - \frac{1}{2}\left(\bar{z}\phi\sigma_I\right)^2$, $\psi_{r^p,z} \equiv \bar{z}(\overline{\hat{\mu}}^R - r - \gamma\sigma_A^2) - \left(\bar{z}\phi\sigma_I\right)^2$, $\psi_{r^p,k} \equiv \bar{z}\psi_{\hat{\mu}^R,k}$.

Wealth dynamics. The log of total wealth for entrepreneur i evolves according to

$$d\log\omega_{i,t} = \left[r^p_{i,t} - \frac{c_{i,t}}{\omega_{i,t}}\right]dt + z_{i,t}\phi\sigma_I dZ_{i,t}, \tag{A.82}$$

Log-linearizing the law of motion of $\omega_{i,t}$, we obtain

$$d\hat{\omega}_{i,t} = \left[\psi_{\omega,0} + \psi_{\omega,\omega}\hat{\omega}_{i,t}\right]dt + \phi\sigma_I z_{i,t}dZ_{i,t}, \tag{A.83}$$

where $\psi_{\omega,0} = \bar{r}^p - \frac{\bar{c}}{\bar{\omega}}$, and $\psi_{\omega,\omega} = \psi_{r^p,z}(\psi_{k,\omega} - 1) + \psi_{r^p,k}\psi_{k,\omega} - \frac{\bar{c}}{\bar{\omega}}(\psi_{c,\omega} - 1)$.

Risk taking. The first-order condition for capital and consumption are given by

$$-V_{\omega\omega}\sigma^{id}_{i,t} = V_\omega p^{id}_{i,t}, \qquad V_\omega = c_{i,t}^{-\gamma}. \tag{A.84}$$

Using the fact that $dV_\omega dZ_{i,t} = V_{\omega\omega}\sigma_{i,t}^{id}dt$ and the expressions above, we can express the optimality condition for capital as follows:

$$p_{i,t}^{id}dt = \gamma \frac{dc_{i,t}}{c_{i,t}} dZ_{i,t}. \tag{A.85}$$

Up to first order, we have that $\frac{dc_{i,t}}{c_{i,t}} dZ_{i,t} = d\hat{c}_{i,t} dZ_{i,t} = \psi_{c,\omega} d\hat{\omega}_{i,t} dZ_{i,t}$. We can then write the expression above as follows:

$$\mu_{i,t}^R - r - \gamma\sigma_A^2 = \gamma\psi_{c,\omega}(\phi\sigma_I)^2 z_{i,t}. \tag{A.86}$$

Evaluating the expression above at $\omega_{i,t} = \overline{\omega}$, we obtain

$$\frac{q\bar{k}}{\overline{\omega}} = \frac{\overline{\mu}^R - r - \gamma\sigma_A^2}{\gamma\psi_{c,\omega}(\phi\sigma_I)^2} = \frac{\overline{p}^{id}}{\gamma\psi_{c,\omega}\phi\sigma_I}. \tag{A.87}$$

Linearizing the optimality condition for capital and matching coefficients, we obtain

$$-\psi_{\mu^R,k}\psi_{k,\omega} = (\overline{\mu}^R - r - \gamma\sigma_A^2)(\psi_{k,\omega} - 1). \tag{A.88}$$

Rearranging the expression above, we obtain

$$\psi_{k,\omega} \equiv \frac{\overline{\mu}^R - r - \gamma\sigma_A^2}{\overline{\mu}^R - r - \gamma\sigma_A^2 + \psi_{\mu^R,k}}. \tag{A.89}$$

Consumption. The envelope condition with respect to ω is given by

$$\hat{\rho}V_\omega = V_\omega \hat{r} + \frac{\mathbb{E}\left[dV_\omega\right]}{dt}. \tag{A.90}$$

Using Ito's lemma, we can write the expression above as follows

$$r = \rho + \gamma\left(\mu_A + \frac{1}{dt}\mathbb{E}[d\hat{c}_{i,t}]\right) - \frac{\gamma(\gamma+1)}{2}\sigma_A^2 - \frac{\gamma^2}{2}\mathbb{E}[(d\hat{c}_{i,t})^2]. \tag{A.91}$$

Up to first order, the expression above can be written as

$$r = \rho + \gamma\left[\mu_A + \psi_{c,\omega}\left(\psi_{\omega,0} + \psi_{\omega,\omega}\hat{\omega}_{i,t}\right)\right] - \frac{\gamma(\gamma+1)}{2}\sigma_A^2 - \frac{\gamma^2}{2}\psi_{c,\omega}^2(\phi\sigma_I)^2 z_{i,t}^2. \tag{A.92}$$

By matching coefficients, we obtain the condition

$$\psi_{r^p,z}(\psi_{k,\omega}-1)+\psi_{r^p,k}\psi_{k,\omega}-\frac{\overline{c}}{\overline{\omega}}(\psi_{c,\omega}-1)=\frac{(\overline{p}^{id})^2}{\gamma\psi_{c,\omega}}(\psi_{k,\omega}-1) \tag{A.93}$$

Rearranging the expression above, we obtain

$$\frac{\overline{c}}{\overline{\omega}}\psi_{c,\omega}^2-\psi_{c,\omega}\left[\psi_{r^p,z}(\psi_{k,\omega}-1)+\psi_{r^p,k}\psi_{k,\omega}+\frac{\overline{c}}{\overline{\omega}}\right]-\frac{(\overline{p}^{id})^2}{\gamma}(1-\psi_{k,\omega}). \tag{A.94}$$

Solving the quadratic equation, we obtain

$$\psi_{c,\omega}=\frac{\overline{\omega}}{2\overline{c}}\left[\psi_{r^p,z}(\psi_{k,\omega}-1)+\psi_{r^p,k}\psi_{k,\omega}+\frac{\overline{c}}{\overline{\omega}}+\sqrt{\left(\psi_{r^p,z}(\psi_{k,\omega}-1)+\psi_{r^p,k}\psi_{k,\omega}+\frac{\overline{c}}{\overline{\omega}}\right)^2+4\frac{(\overline{p}^{id})^2}{\gamma}(1-\psi_{k,\omega})\frac{\overline{c}}{\overline{\omega}}}\right]. \tag{A.95}$$

It remains to determine the intercept:

$$r=\rho+\gamma\left[\mu_A+\psi_{c,\omega}\left(\overline{r}^p-\frac{\overline{c}}{\overline{\omega}}\right)\right]-\frac{\gamma(\gamma+1)}{2}\sigma_A^2-\frac{\gamma^2}{2}\psi_{c,\omega}^2(\phi\sigma_I)^2\overline{z}^2. \tag{A.96}$$

We can solve for the consumption-wealth ratio:

$$\frac{\overline{c}}{\overline{\omega}}=\frac{\rho}{\gamma\psi_{c,\omega}}+\left(1-\frac{1}{\gamma\psi_{c,\omega}}\right)\left[r+\frac{(p^{ag})^2}{2\gamma}+\frac{(\overline{p}^{id})^2}{2\gamma\psi_{c,\omega}}\right]+\left[\frac{\gamma\sigma_A^2}{2}-\mu_A\right]\left(1-\psi_{c,\omega}^{-1}\right)+\overline{z}(\overline{\hat{\mu}}^R-\overline{\mu}^R). \tag{A.97}$$

The expression above reduces to $\frac{\overline{c}}{\overline{\omega}}=\overline{r}$ in the case with constant returns to scale.

Determining the approximation point. At the point $\hat{\omega}_{i,t}=0$, the drift of $\hat{\omega}_{i,t}$ is zero, which implies that $\overline{r}^p=\frac{\overline{c}}{\overline{\omega}}$. The triplet $(\overline{c}_j,\overline{k}_j,\overline{\omega}_j)$ satisfies the conditions:

$$r=\rho+\gamma\mu_A-\frac{\gamma(\gamma+1)}{2}\sigma_A^2-\frac{(\overline{p}_j^{id})^2}{2},\qquad \overline{z}_j=\frac{\overline{p}^{id}}{\gamma\psi_{c,\omega}\phi\sigma_I^j},\qquad \overline{r}_j^p=\frac{\overline{c}_j}{\overline{\omega}_j}, \tag{A.98}$$

where $\overline{r}^p=r+\gamma\sigma_A^2-\mu_A+\frac{(\overline{p}^{id})^2}{\gamma\psi_{c,\omega}}-\frac{1}{2}\left(\frac{\overline{p}^{id}}{\gamma\psi_{c,\omega}}\right)^2+\overline{z}(\overline{\hat{\mu}}^R-\overline{\mu}^R)$. These equations give the value of $(\overline{c},\overline{k},\overline{\omega})$ given $\psi_{c,\omega}$. We can then iterate between the computation of $\psi_{c,\omega}$ and $(\overline{c},\overline{k},\overline{\omega})$ until convergence. Using the fact that $r=\rho_w+\gamma\mu_A-\frac{\gamma(\gamma+1)}{2}\sigma_A^2$, we can solve for the price of idiosyncratic risk:

$$\overline{p}_j^{id}=\sqrt{2(\rho-\rho_w)}. \tag{A.99}$$

Hence, the unconditional price of idiosyncratic risk is equalized across entrepreneurs.

Capital for type j satsifies the condition:

$$r + p^{ag}\sigma_A + \overline{p}^{id}_j \phi \sigma_I^j = \frac{\alpha K^{-\frac{\alpha\beta}{1-\beta}} \overline{k}_j^{\frac{\alpha}{1-\beta}-1} e_i^{\frac{1-\alpha-\beta}{1-\beta}} - \iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta. \tag{A.100}$$

We can rearrange the expression above as follows:

$$\zeta_j = K^{-\frac{\alpha\beta}{1-\beta}} \overline{k}_j^{\frac{\alpha+\beta-1}{1-\beta}} e_j^{\frac{1-\alpha-\beta}{1-\beta}} \Rightarrow \sum_{j=1}^{n} \tilde{\chi}_{e,j} e_j \zeta_j^{\frac{\alpha}{\alpha+\beta-1}} = K^{-\frac{\alpha\beta}{1-\beta}\frac{\alpha}{\alpha+\beta-1}} \sum_{j=1}^{n} \tilde{\chi}_{e,j} \overline{k}_j^{\frac{\alpha}{1-\beta}} e_j^{\frac{1-\alpha-\beta}{1-\beta}} \tag{A.101}$$

where $\zeta_j \equiv \alpha^{-1}\left\{q\left[r + p^{ag}\sigma_A + \overline{p}^{id}\phi\sigma_I^j - (\mu_A + \Phi(\iota(q)) - \delta)\right] + \iota(q)\right\}$.

Rearranging the expression above, we obtain

$$\zeta = K^{\alpha-1}, \tag{A.102}$$

where $\zeta = \left[\sum_{j=1}^{n} \tilde{\chi}_{e,j} e_j \zeta_j^{\frac{\alpha}{\alpha+\beta-1}}\right]^{\frac{\alpha+\beta-1}{\alpha}}$. The value of $\overline{k}_j$ is then given by $\overline{k}_j = e_j K \left(\frac{\zeta_j}{\zeta}\right)^{\frac{1-\beta}{\alpha+\beta-1}}$.

Price of idiosyncratic risk. The price of idiosyncratic risk is given by

$$p^{id}_{i,t} = \frac{\mu^R_{i,t} - r - \gamma\sigma_A^2}{\phi\sigma_{I,i}} = \overline{p}^{id} - \frac{\psi^j_{\mu^R,k}}{\phi\sigma_I^j}\psi^j_{k,\omega}\hat{\omega}_{i,t}. \tag{A.103}$$

Using the fact that $\mathbb{E}[\hat{\omega}_{i,t}]$ by definition of the approximation point, we obtain $\mathbb{E}[p^{id}_{i,t}] = \overline{p}^{id}$. Unconditional expected marginal returns are then given by

$$\mathbb{E}[\mu^R_{i,t}] = r + p^{ag}\sigma_A + \overline{p}^{id}\phi\sigma_{I,i}. \tag{A.104}$$

□

A.7 Proof of Proposition 6

Proof. Aggregating Equation (7) and using the fact that labor supply grows at rate g, we obtain

$$\dot{K}_t = \left[\Phi(\iota(q_t)) - \delta - g\right] K_t, \tag{A.105}$$

given the initial condition $K_0 = K^*$. From (A.15), we obtain the expression

$$r + p_t^{ag}\sigma_A + p_t^{id}\phi\sigma_I = \frac{\alpha K_t^{\alpha-1} - \iota(q_t)}{q_t} + \frac{\dot{q}_t}{q_t} + \Phi(\iota(q_t)) - \delta + \mu_A. \tag{A.106}$$

Using $p_t^{ag} = \gamma\sigma_A$ and $p^{id} = \gamma\phi\sigma_I \frac{q_t K_t}{\chi_e \omega_{e,t}}$ and solving for $\dot{q}_t$, we obtain

$$\dot{q}_t = \left[r + \gamma\sigma_A^2 + \gamma\phi^2\sigma_I^2 \frac{q_t K_t}{\chi_e \omega_{e,t}} + \delta - \mu_A - \Phi(\iota(q_t)) - \frac{\alpha K_t^{\alpha-1} - \iota(q_t)}{q_t} \right] q_t, \tag{A.107}$$

where $\omega_{e,t} = n_{e,t} + h_{e,t}$.

The ODE above is subject to the terminal condition

$$\lim_{t\to\infty} q_t = q, \tag{A.108}$$

where q is the value of q_t in the new stationary equilibrium.

The first PDE was derived in the proof of Lemma 1 and it was given in (A.18). The boundary conditions are $\zeta_t(T) = (1-\psi)^{-1} (V^*)^{-\frac{1}{\gamma}}$ and

$$\lim_{t\to\infty} \zeta_t(a) = \zeta(a), \tag{A.109}$$

where $\zeta(a)$ is the value in the new stationary equilibrium.

The PDE for the human wealth was given in (A.7). The boundary conditions are $h_t(T) = 0$ and

$$\lim_{t\to\infty} h_t(a) = h(a), \tag{A.110}$$

where $h(a)$ is the value in the new stationary equilibrium.

The PDE for total wealth was derived in the proof of Proposition 2 and it was given in (A.32). The first boundary condition is $\omega_t(0) = e^{-gT}\omega_t(T) + h_t(0)$. The initial condition for $\omega_t(a)$ is given by

$$\omega_0(a) = n^*(a) + (q_0 - q^*)k^*(a) + h_0(a), \tag{A.111}$$

where variables with an asterisk denote values before the change in ϕ.

□

B Data

We next discuss our empirical measures in more detail. For an extensive description of the Townsend Thai Monthly Survey (TTMS) and the derivation of entrepreneurs' balance sheet information from the survey questionnaire, see Samphantharak and Townsend (2010).

B.1 Sample selection and variable definition

The dataset contains both economic and demographic variables. Economic variables include households' assets and liabilities, financial wealth, business and labor income, and consumption. Demographic and geographic variables consist of the age of the household head, year, number of household members, number of children, and province. We restrict the sample to households with heads between 25 and 80 years old. Observations without information on age or financial wealth are excluded, resulting in an unbalanced panel of 710 households over 14 years, from 1999 to 2012. The subset with consistent information on both assets and income, which is required for return calculations, includes 541 households.

Business exposure. The TTMS reports detailed information on entrepreneurs' assets, including fixed assets, inventories, and financial assets. We classify these into business and non-business (safe) assets. The value of the business includes inventories, livestock, agricultural assets, business assets, and household assets. Following Samphantharak and Townsend (2018), we include household assets such as cars, pick-up trucks, or fishing boats, since these are frequently used in production activities. Safe assets consist of cash in hand, account receivables, deposits at financial institutions, ROSCA, other lending, prepaid insurance, and land. We then compute the fraction of financial wealth (total assets net of liabilities) invested in the business, which serves as our measure of risk-taking.

Return on business activity. Given estimates of business value and the flow of business income generated in a period, we compute the return on assets (ROA) as the ratio of business income to the value of business assets. ROA is a standard accounting measure of profitability. Our measure includes realized capital gains, which are observed in the survey, but excludes unrealized capital gains. In Section 5, we show that differences in expected returns in our theory are primarily driven by dividend yields rather than unrealized capital gains, so theoretical returns are broadly consistent with how returns are measured empirically. Because ROA may still be subject to measurement error, we

Table B.1: Descriptive Statistics of Households Characteristics

Variable	Q1	Median	Mean	Q3	Count
Age	43.00	53.00	53.49	64.00	8858
Household size	3.00	4.00	3.86	5.00	7846
Number of children	0.00	1.00	0.94	2.00	7846
Consumption-wealth ratio	0.03	0.07	0.12	0.14	8858
Business exposure	0.09	0.19	0.24	0.35	8858
Human-financial wealth ratio	0.11	0.30	0.76	0.67	8858
Net worth (normalized)	0.18	0.38	1.00	0.84	8858

Note: The columns $Q1$ and $Q3$ refer to the first and third quartiles of the distribution, respectively. Business exposure is the ratio of the value of the business to the net worth. Human-financial wealth ratio is the ratio of human wealth to financial wealth. The normalized net worth for each household corresponds to the ratio of their net worth to the average net worth in that year.

provide several robustness checks: the portfolio-level analysis in Section 2, the use of shrinkage estimators in Appendix B.3, and an explicit treatment of measurement error in the latent factor model in Appendix B.5.

Human wealth. We construct our empirical measure of human wealth analogously to its counterpart in the model, as the present discounted value of future expected labor income. This requires specifying a discount rate and an expectation of future labor income. We use the same discount rate as in the model (see Section 3). Expected labor income is age-dependent and computed as the average labor income of households at each age. As in the construction of life-cycle profiles, we use trimmed means to limit the influence of outliers. Following the model, we abstract from idiosyncratic variation in labor income and treat human wealth as purely age-dependent, corresponding to the present discounted value of future average income at each age.

B.2 Descriptive statistics

We next present descriptive statistics for the main variables used in the analysis. Table B.1 reports household-level demographic characteristics, such as age, household size, and number of children, together with economic variables including the consumption-wealth ratio, business exposure (the ratio of business value to net worth), the human-financial wealth ratio, and net worth normalized by the average net worth in that year. Normalization by year helps abstract from aggregate shocks and makes distributions comparable over time.

Table B.2 reports the percentage of revenue from each activity by province. This breakdown highlights the sectoral and geographic heterogeneity across the four provinces in our sample, spanning semi-urban regions near Bangkok and rural areas in the northeast.

Table B.2: Production Activity by Province

Region:	Central		Northeast	
Province:	**Chachoengsao**	**Lopburi**	**Buriram**	**Srisaket**
Production activity (%):				
Cultivation	13.30	44.10	16.90	40.30
Livestock	27.70	21.90	1.10	1.20
Fish and shrimp	18.60	0.00	0.30	1.80
Nonfarm business	27.40	23.40	62.90	31.50
Wage earning	12.90	10.50	18.80	25.20

Such heterogeneity provides important variation in business opportunities and risks that we exploit in our empirical analysis.

B.3 Shrinkage estimators

To perform the cross-sectional regression in the second stage, we must first estimate the idiosyncratic variance, the aggregate factor loadings, and the individual return means for each entrepreneur. Because the number of parameters is large relative to the time dimension, individual estimates may be noisy, raising concerns of *overfitting*. We therefore consider a range of shrinkage estimators that reduce estimation noise by pulling individual estimates toward the cross-sectional mean.

Ledoit-Wolf shrinkage estimator. We begin with the covariance matrix of (residual) returns. Let $\hat{\epsilon}_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_t^{agg}$ denote the residuals from the first-stage regressions for $i \in \{1, \ldots, N\}$, and let $\hat{\Sigma}$ be the $N \times N$ sample covariance matrix of residuals. The diagonal elements of $\hat{\Sigma}$ are the idiosyncratic variances, $\hat{\sigma}_i^2$. The Ledoit-Wolf shrinkage estimator is

$$\hat{\Sigma}_{LW} = (1 - \rho_{LW})\hat{\Sigma} + \rho_{LW}\overline{\sigma}^2 I_N, \tag{B.1}$$

where ρ_{LW} is the shrinkage coefficient and $\overline{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} \hat{\sigma}_i^2$ is the average variance. The diagonal elements of $\hat{\Sigma}_{LW}$ provide shrinkage-adjusted idiosyncratic variances, $\hat{\sigma}_{i,LW}^2$. We compute ρ_{LW} using the Ledoit-Wolf formula.

James-Stein estimator for the mean. An analogous approach applies to mean returns. The James-Stein estimator is

$$\hat{\mu}_{JS} = (1 - \rho_{JS})\hat{\mu} + \rho_{JS}\overline{\mu}, \tag{B.2}$$

Table B.3: Cross-sectional regressions with shrinkage estimators

	Mean ROA		Mean ROA (JS)		Mean ROA		Mean ROA (JS)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Beta	0.012***	0.012***	0.011***	0.011***				
	(0.001)	(0.001)	(0.001)	(0.001)				
Id. Variance	0.103***		0.097***					
	(0.004)		(0.004)					
Id. Variance (LW)		0.221***		0.208***				
		(0.009)		(0.009)				
Beta (RE)					0.017***	0.017***	0.016***	0.016***
					(0.001)	(0.001)	(0.001)	(0.001)
Id. Variance (RE)					0.091***		0.086***	
					(0.004)		(0.004)	
Id. Variance (RE/LW)						0.194***		0.183***
						(0.009)		(0.008)
Observations	541	541	541	541	541	541	541	541
R^2	0.685	0.685	0.685	0.685	0.734	0.734	0.734	0.734
Adj. R^2	0.684	0.684	0.684	0.684	0.733	0.733	0.733	0.733

Note: Id. Variance (LW) refers to the Ledoit-Wolf estimate for the idiosyncratic variance. Id. Variance (RE) refers to the idiosyncratic variance of the residuals from the random effects regression. Id. Variance (RE/LW) refers to the Ledoit-Wolf estimate for the idiosyncratic variance of the residuals from the random effects regression. Mean ROA (JS) refers to the James-Stein estimate for the mean.

where $\hat{\mu} = [\hat{\mu}_1, \ldots, \hat{\mu}_N]^\top$ is the vector of sample means, $\overline{\mu} = \frac{1}{N}\sum_{i=1}^N \hat{\mu}_i$ is the cross-sectional average, and ρ_{JS} is the James-Stein shrinkage coefficient.

Shrinkage estimator for factor loadings. Finally, we apply shrinkage to factor loadings using a mixed-effects model. The first-stage regression can be written as

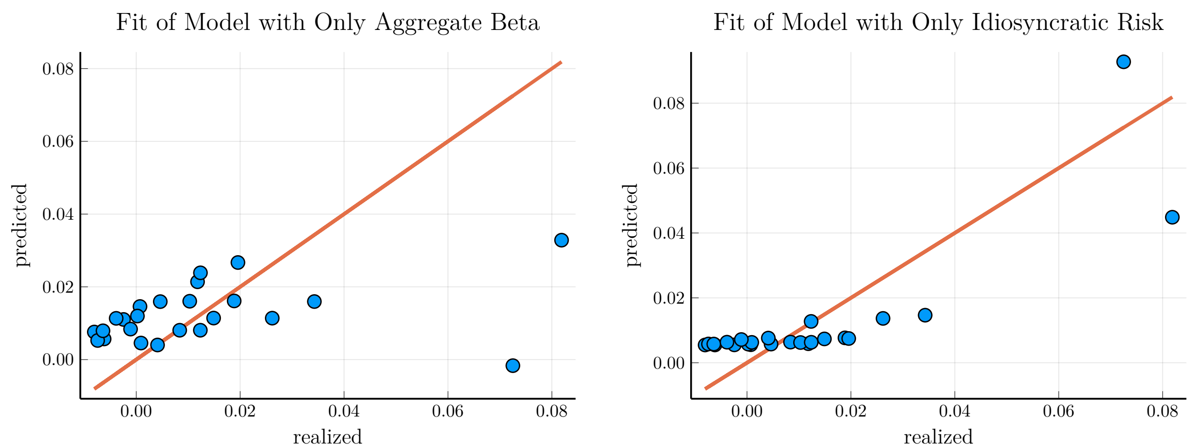
$$R_{i,t} \mid \beta_i, R_t^{agg} \sim \mathcal{N}\left(\alpha_i + \beta_i R_t^{agg}, \sigma_i^2\right), \tag{B.3}$$

where the vector of factor loadings $\beta = [\beta_1, \ldots, \beta_N]^\top$ follows $\beta \sim \mathcal{N}(\overline{\beta}\mathbf{1}_N, \Sigma_\beta)$. This specification shrinks the individual slope coefficients β_i toward the cross-sectional average $\overline{\beta}$. We estimate the model by maximum likelihood, which delivers shrinkage-adjusted estimates of factor loadings that are less sensitive to sampling variation.

Cross-sectional regressions. The shrinkage methods described above mitigate potential overfitting in the first-stage estimates. Given these adjusted parameters, we estimate the second-stage cross-sectional regressions reported in Table B.3.

Column 1 reproduces the baseline results from column 3 of Table 1 for ease of comparison. Columns 2, 4, 6, and 8 apply the Ledoit–Wolf estimator to the idiosyncratic variance, while columns 3, 4, 7, and 8 use the James–Stein shrinkage for the mean re-

Figure B.1: Realized vs. predicted returns: one-factor models



Note: The left (right) panel shows a scatter plot of predicted returns of the single-factor model with aggregate beta (idiosyncratic variance) as the sole factor against realized returns in the portfolio-level analysis.

turn. Consistent with the presence of attenuation bias in the coefficient on idiosyncratic variance, the shrinkage corrections raise this coefficient substantially—roughly doubling its magnitude relative to the baseline. The resulting estimates are in line with those obtained in the portfolio-level analysis, which is less sensitive to measurement error (see column 4 of Table 1). Shrinking the mean return using the James–Stein estimator has only a minor quantitative effect, as shown by the small differences between columns 3 and 4 (and between 7 and 8). When factor loadings are estimated using the random-effects model (columns 5–8), the coefficient on aggregate beta is somewhat larger and that on idiosyncratic variance somewhat smaller, but both remain within the same range as in the portfolio regressions. Overall, the similarity of estimates across all specifications indicates that our main results are robust to overfitting and noise in the first-stage estimates.

B.4 Fit of single-factor models

Figure 2 in Section 2 illustrates the performance of the two-factor model. Here we assess the fit of single-factor models that rely only on aggregate beta or only on idiosyncratic variance. Figure B.1 shows that both one-factor specifications systematically underfit the dispersion in entrepreneurial returns, highlighting the importance of considering both sources of risk jointly. This visual evidence is consistent with the regression results in Table 1, where both factors are required to explain the cross-section of returns.

B.5 Model with latent factors

We consider next the model with latent factors presented in Section 2.1 in more detail. First, we show how to accommodate time-varying factor loadings and measurement error in the model. Second, we discuss the estimation and construction of standard errors in more detail, including how we handled missing values.

Time-varying loadings. One concern with our baseline specification is that factor exposures may vary with aggregate conditions. Consider a version of the latent factor model with time-varying loadings:

$$R_{i,t+1} = \mu_{i,t} + \beta_{i,t}^{\top} f_{t+1} + \epsilon_{i,t+1}. \tag{B.4}$$

Following, e.g., Lettau and Ludvigson (2001), we assume that loadings depend on an aggregate variable $z_t \in \mathbb{R}$:

$$\beta_{i,t} = \eta_{i,0} + \eta_{i,z} z_t, \qquad \mathbb{E}[z_t] = 0. \tag{B.5}$$

Conditional expected returns are then

$$\mu_{i,t} = \lambda_0 + \lambda_{ag}^{\top} \beta_{i,t} + \lambda_{id} \sigma_i^2. \tag{B.6}$$

Substituting yields a representation with an extended factor vector $\tilde{f}_t \equiv [f_t^{\top}, (z_t f_{t+1} - \mathbb{E}[z_t f_{t+1}])^{\top}, z_t]^{\top}$, factor loadings $\tilde{\beta}_i \equiv [\eta_{i,0}^{\top}, \eta_{i,z}^{\top}, \lambda_{ag}^{\top} \eta_{i,z}]^{\top}$, and risk prices $\tilde{\lambda}_{ag} \equiv [\lambda_{ag}^{\top}, \mathbb{E}[z_t f_{t+1}]^{\top}, 0]^{\top}$. The model can then be written in the familiar form

$$R_{i,t} = \mu_i + \tilde{\beta}_i^{\top} \tilde{f}_t + \epsilon_{i,t}, \tag{B.7}$$

where $\mu_i = \lambda_0 + \tilde{\lambda}_{ag}^{\top} \tilde{\beta}_i + \lambda_{id} \sigma_i^2$.

Hence, the latent factor framework naturally accommodates time-varying loadings and, by extension, time-varying risk premia.

Measurement error. To address the fact that survey-based measures of business returns are potentially subject to measurement error, we explicitly incorporate measurement error into the latent factor model. We assume that observed returns are given by $\tilde{R}_{i,t} = R_{i,t} + v_{i,t}$ with $v_{i,t} \sim \mathcal{N}(0, \sigma_v^2)$. The error $v_{i,t}$ is independent of the idiosyncratic shock $\epsilon_{i,t}$ and the aggregate factors f_t. The residual variance then combines both idiosyncratic shocks and measurement error: $\tilde{\sigma}_i^2 = \sigma_i^2 + \sigma_v^2$. Expected returns, however, reflect only the compensation for idiosyncratic shocks:

$$\mu_i = \lambda_{ag}^\top \beta_i + \lambda_{id} \sigma_i^2. \tag{B.8}$$

Observed returns can therefore be written as

$$\tilde{R}_{i,t} = \tilde{\lambda}_0 + \lambda_{ag}^\top \beta_i + \lambda_{id} \tilde{\sigma}_i^2 + \beta_i^\top f_t + \tilde{\epsilon}_{i,t}, \tag{B.9}$$

where $\tilde{\lambda}_0 = -\lambda_{id}\sigma_v^2$. Thus, classical measurement error shifts only the intercept of the pricing equation but not the estimated prices of risk. This provides a natural explanation for the negative, and often insignificant, intercepts in Table 1, while leaving our estimates of risk premia unaffected.

Estimation. We estimate the latent-factor model using the three-pass estimator of Giglio et al. (2021). In the first stage, we run time-series regressions of returns on the observable factor:

$$R_{i,t} = \mu_i + \beta_{i,o}^\top f_{o,t} + \tilde{\epsilon}_{i,t}, \tag{B.10}$$

where $\tilde{\epsilon}_{i,t} = \beta_{i,u}^\top f_{u,t} + \epsilon_{i,t}$. In the second stage, we apply PCA to the first-stage residuals to estimate $\beta_{i,u}$, the loadings on the latent factors. In the third stage, we regress average returns on the observable loadings $\hat{\beta}_{i,o}$, the latent loadings $\hat{\beta}_{i,u}$, and the variance of the PCA residuals. We do not use the variance of the first-stage residuals, since those include both idiosyncratic shocks and latent-factor exposures.

Importantly, $\hat{\beta}_{i,o}$ may not consistently estimate $\beta_{i,o}$ if observable factors are correlated with latent ones. However, Giglio et al. (2021) show that $(\hat{\beta}_{i,o}, \hat{\beta}_{i,u})$ converge in probability to a rotation of the true loadings, $H\beta_i$, where H is a rotation matrix. As a result, the aggregate risk premium $\lambda_{ag}^\top \beta_i = (H\lambda_{ag})^\top (H\beta_i)$ is consistently estimated, since it is invariant to the rotation.

Missing values. Standard PCA methods require complete data. Given our unbalanced panel, we adopt an iterative procedure to compute the principal components. PCA can be obtained from the singular value decomposition (SVD) of the $N \times T$ matrix of returns R, which in our case has missing entries. To initialize, we impute missing values using a simple method such as the sample mean return. We then apply the SVD to the completed matrix, obtain a low-rank approximation based on the first k singular values, and update the imputed values using this approximation. We repeat this procedure until convergence. Giglio et al. (2021) show that the three-pass estimator remains consistent when

PCA requires such a matrix completion step.[32]

Standard errors. Following Giglio et al. (2021), we compute standard errors for the three-pass estimator using a wild bootstrap. Formally, we generate B bootstrap samples by resampling the residuals from the PCA step with random weights that satisfy zero mean and unit variance (see Mammen, 1993). For each bootstrap sample, we re-estimate the factor loadings and risk prices, obtaining B estimates of $(\hat{\lambda}_0^b, \hat{\lambda}_{ag}^b, \hat{\lambda}_{id}^b)$. The standard errors are given by the standard deviation of these estimates across bootstrap simulations.

Testing the number of latent factors. We test the number of latent factors using the procedure proposed by Onatski (2009). Specifically, we test the null hypothesis that the residuals of the first-stage regression contain no latent factors. The test statistic is -0.32, while the 5% critical value from the Tracy–Widom distribution is 1.27. Since the statistic falls well below the critical value, we cannot reject the null hypothesis that the number of latent factors is zero. This result reinforces the view that the observable factor captures nearly all the common variation in entrepreneurial returns.

Testing the risk-based model. We also assess whether the risk-based model—where differences in expected returns are fully explained by exposure to aggregate and idiosyncratic risk—is consistent with the data. Expected returns for entrepreneur i are given by

$$\mu_i = \lambda_0 + \lambda_{ag}^{\top}\beta_{i,o} + \lambda_{id}\sigma_i^2 + u_i, \tag{B.11}$$

where u_i captures any non-risk-based component. We test the joint hypothesis that $u_i = 0$ for all i, computing the bootstrap p-value for the test statistic $T_u = \sum_{i=1}^{N} u_i^2$. To limit the influence of measurement error, we conduct the test at the portfolio level. For the model with latent factors, the p-value is 0.082, and for the model with only observable factors, it is 0.214. In both cases, we cannot reject the null hypothesis at the 5% level, suggesting that risk exposures account for most of the variation in expected returns.

Taken together, these results support the main empirical finding in Section 2.1: differences in expected entrepreneurial returns are well explained by exposure to aggregate and idiosyncratic risk, with little evidence of omitted latent factors or non-risk-based components.

[32]They also show that a correction is necessary for hypothesis testing on the residuals of the cross-sectional regression in the presence of missing values. In our case, only about 1.5% of observations are missing, so we found this correction to be negligible.

B.6 Life-cycle profiles

To construct life-cycle profiles for a given variable, we aggregate households into age groups, with cutoffs chosen so that each group contains roughly the same number of households. While the main text reports results based on five broad age groups, here we consider a more flexible specification with 15 groups as a robustness check. The trade-off is that with more groups, each cell contains fewer observations, which can make the estimates noisier. To mitigate this issue and reduce the influence of outliers, we compute trimmed means using a 7.5% trimming on each side, which is more aggressive than in the main text and helps stabilize the estimates with a larger number of groups.

Let $z_{i,k,t}$ denote variable z for household i in age group k at year t. We estimate the following specification:

$$z_{i,k,t} = \alpha_t + age_k + \delta' x_{i,k,t} + u_{i,k,t}, \tag{B.12}$$

where α_t represents year fixed effects, age_k denotes the age-group effect, and $x_{i,k,t}$ is a vector of demographic and geographic controls (household size, number of children, province dummies, and sector dummies).

The raw age-group effect is given by

$$age_k^{raw} = \mathbb{E}\left[z_{i,k',t} \mid k' = k\right], \tag{B.13}$$

which can be estimated either by simple age-group averages or by regressing $z_{i,k,t}$ on a set of age-group dummies.

We follow Kaplan (2012) and define the age-group effect controlling for year fixed effects as

$$age_k^{year-FE} = \frac{1}{\overline{T}} \sum_{t=1}^{\overline{T}} \mathbb{E}\left[z_{i,k',t'} \,\middle|\, k' = k,\, t' = t\right]. \tag{B.14}$$

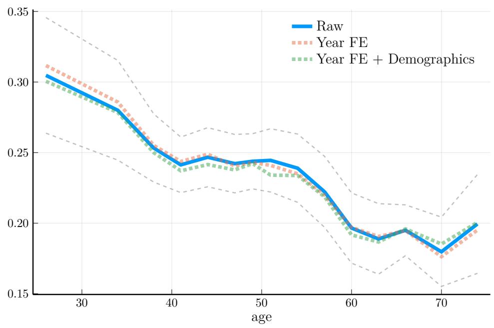
This can be estimated by regressing $z_{i,k,t}$ on dummies for age groups and year fixed effects, and then computing the predicted value at age k evaluated at an "average" year.

Similarly, we define the age-group effect controlling for both year fixed effects and demographic/geographic controls as

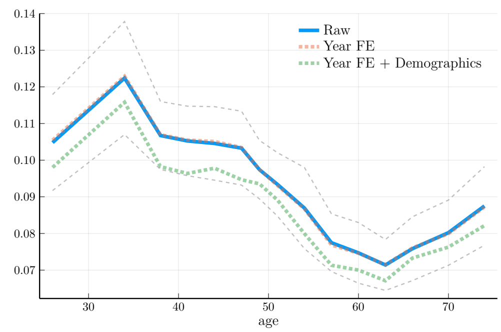
$$age_k^{year-FE+dem} = \frac{1}{\overline{T}} \sum_{t=1}^{\overline{T}} \mathbb{E}\left[z_{i,k',t'} \,\middle|\, k' = k,\, t' = t,\, x_{i,k',t'} = \overline{x}_{k,t}\right], \tag{B.15}$$

where $\overline{x}_{k,t}$ denotes the average value of controls $x_{i,k,t}$ for age group k in year t. This can be estimated from the full regression with year and age-group fixed effects as well as demographic controls.

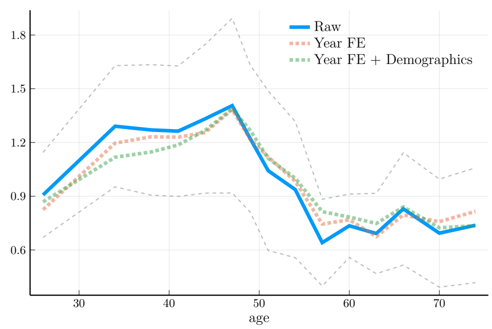
Figure B.2: Life-cycle profiles: raw and additional controls



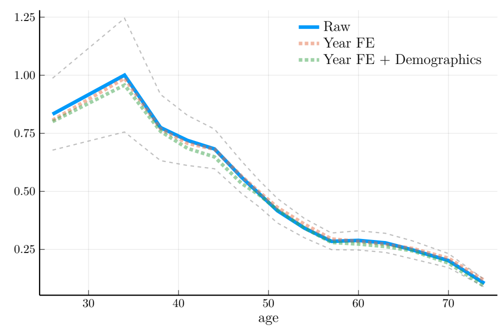
(a) Business exposure



(b) Consumption-wealth ratio



(c) Labor income



(d) Human-financial wealth ratio

Figure B.2 plots the estimated age-group effects for the main variables. Grey dashed lines denote the 95% confidence interval for the raw estimates. Standard errors are clustered by household and year. The life-cycle patterns based on the raw estimates are very similar to those obtained after including year fixed effects and demographic controls, and the differences are not statistically significant.

Table B.4 reports formal tests of equality across age-group coefficients. For the risk-taking measure, we reject the null that the youngest and oldest groups have the same coefficient, consistent with the observed decline in risk-taking over the life cycle. For the consumption-wealth ratio, we reject equality between the youngest group and group 12, which has the lowest coefficient, as well as between group 12 and the oldest group. Finally, we reject equality between the youngest and oldest groups. Together, these results confirm the U-shaped pattern of the consumption-wealth ratio: it declines from young to middle age, before increasing again at older ages. Overall, the tests in Table B.4 formally validate the descriptive patterns highlighted in Section 2.

Overall, the results confirm that the life-cycle patterns highlighted in Section 2 are robust both to the inclusion of additional controls and to adopting a more flexible age-group specification.

Table B.4: Formal tests of equality of life-cycle coefficients

Variable	Hypothesis tested	p-value
Risk-taking	H_0: age_group01 = age_group15	3.3×10^{-4}
Consumption-wealth ratio	H_0: age_group01 = age_group12	1.1×10^{-5}
Consumption-wealth ratio	H_0: age_group12 = age_group15	9.7×10^{-3}
Consumption-wealth ratio	H_0: age_group01 = age_group15	3.9×10^{-2}

Note: We test whether life-cycle coefficients on age dummies are statistically different. For the risk-taking measure, we test equality between the youngest and oldest age groups (1 and 15). For the consumption-wealth ratio, we test equality between the youngest group and the group with the lowest coefficient (age_group12), between that minimum and the oldest group, and between the youngest and the oldest groups. All tests are based on the specification without controls. p-values are obtained from Wald tests with clustered standard errors (household and year).

C Derivations

C.1 Equilibrium prices and capital stock in a stationary equilibrium

Interest rate. The financial wealth of wage earners evolves according to

$$d\tilde{n}_{j,t} = \left[(r + \gamma\sigma_A^2)\tilde{n}_{j,t} + \tilde{w}_t \bar{l}_{j,t} - \tilde{c}_{j,t}\right] dt + \tilde{n}_{j,t}\sigma_A dZ_t. \tag{C.1}$$

using the demand for aggregate insurance $\theta_{j,t} = -n_{j,t}\sigma_A$.

Combining the expression above with the law of motion for human wealth, we obtain the law of motion of total wealth:

$$\frac{d\tilde{\omega}_{i,t}}{\tilde{\omega}_{i,t}} = \left[r + \gamma\sigma_A^2 - \frac{c_{j,t}}{n_{j,t} + h_{j,t}}\right] dt + \sigma_A dZ_t. \tag{C.2}$$

In a stationary equilibrium, wage earners' scaled total wealth, $\omega_{j,t} = \tilde{\omega}_{j,t}/A_t$, is constant. Therefore, the drift of $\omega_{j,t}$ must be zero. Using Ito's lemma, we obtain:

$$\frac{d\omega_{i,t}}{\omega_{i,t}} = \frac{d\tilde{\omega}_{i,t}}{\tilde{\omega}_{i,t}} - \frac{dA_t}{A_t} + \left(\frac{dA_t}{A_t}\right)^2 - \frac{dA_t}{A_t}\frac{d\tilde{\omega}_{i,t}}{\tilde{\omega}_{i,t}} \tag{C.3}$$

$$= \left[r + \gamma\sigma_A^2 - \frac{c_{j,t}}{n_{j,t} + h_{j,t}} - \mu_A\right] dt. \tag{C.4}$$

The interest rate must then satisfy the condition

$$r + \gamma\sigma_A^2 - \left[\frac{1}{\gamma}\rho_w + \left(1 - \frac{1}{\gamma}\right)\left(r + \frac{\gamma\sigma_A^2}{2}\right)\right] - \mu_A = 0, \tag{C.5}$$

using the fact that the consumption-wealth ratio is given by

$$\frac{c_{j,t}}{n_{j,t}+h_{j,t}} = \frac{1}{\gamma}\rho_w + \left(1-\frac{1}{\gamma}\right)\left(r + \frac{(p^{ag})^2}{2\gamma}\right), \tag{C.6}$$

which is a special case of (24), as we set $p_t^{id} = 0$ and $T \to \infty$.

Rearranging expression (C.5), we obtain

$$r = \rho_w + \gamma\mu_A - \gamma(\gamma+1)\frac{\sigma_A^2}{2}. \tag{C.7}$$

Relative price of capital. Plugging the expression for $\Phi'(\iota)$ into the first-order condition for ι in equation (16), we obtain

$$\frac{1}{\sqrt{\overline{\Phi}^2 + 2\iota}} = \frac{1}{q} \Rightarrow \iota = \frac{q^2 - \overline{\Phi}^2}{2}. \tag{C.8}$$

In a stationary equilibrium, the capital-labor ratio is constant. Thus, capital grows at the population rate g, which gives us the condition

$$\Phi(\iota) - \delta = g \Rightarrow q = \overline{\Phi} + (g+\delta), \tag{C.9}$$

obtained by plugging the expression for ι into the functional form for $\Phi(\iota)$.

C.1.1 Capital stock and idiosyncratic risk premium

Rearranging the expression for the shadow price of idiosyncratic risk (22), and using the definition of expected returns (17), we obtain the MPK schedule

$$r + p^{ag}\sigma_A + p^{id}\phi\sigma_I = \frac{\alpha K^{\alpha-1} - \iota(q)}{q} + \mu_A + \Phi(\iota(q)) - \delta. \tag{C.10}$$

where r, p^{ag}, and q are functions of parameters, as derived above.

Integrating condition (21) across all entrepreneurs, we obtain

$$p^{id} = \gamma\phi\sigma_I\frac{qK}{\chi_e\omega_e}, \tag{C.11}$$

where χ_e is the fraction of entrepreneurs in the population.

C.2 Financial autarky

We consider next the case of financial autarky, where entrepreneurs have no access to either aggregate or idiosyncratic insurance. To shut down idiosyncratic insurance, we must set $\phi = 1$. To capture the absence of aggregate insurance, we will focus on the special case where there is no demand for aggregate insurance, so the solution would coincide with the case where entrepreneurs have no access to insurance.

Suppose that $h_{i,t} = 0$, so entrepreneurs have no labor income, and assume that $n_{j,t} = 0$ for $j \in \mathcal{W}_t$, so wage earners have no financial wealth. The first assumption implies that $\frac{qk_{i,t}}{n_{i,t}}$ is equalized across entrepreneurs, and the second assumption implies that $qk_{i,t} = n_{i,t}$. As the demand for aggregate insurance is given by $\theta^{ag}_{i,t} = (qk_{i,t} - n_{i,t})\sigma_A$, we obtain that entrepreneurs do not demand aggregate insurance. The solution will then coincide with the case where aggregate insurance is unavailable.

Under these assumptions, the price of idiosyncratic risk, given in Equation (35), specializes to $p^{id} = \gamma\sigma_I$. The risk premium is then given by

$$p^{ag}\sigma_A + p^{id}\sigma_I = \gamma\left[\sigma_A^2 + \sigma_I^2\right]. \tag{C.12}$$

The aggregate Sharpe ratio relative to the idiosyncratic Sharpe ratio is $p^{ag}/p^{id} = \sigma_A/\sigma_I$.

D Extensions

In this section, we describe four extensions to the model, where we introduce the following features: i) limited pledgeability of physical capital, ii) endogenous occupational choice, iii) uninsurable labor income shocks, and iv) Epstein-Zin preferences for workers. The first three extensions are meant to show the robustness of the main results to introducing ingredients typically present in models of entrepreneurship. The last extension shows how one can obtain a constant interest rate along the transitional dynamics, which enable us to focus on the implications of changes in the idiosyncratic risk premium.

D.1 Limited pledgeability and heterogeneous expected returns

In this subsection, we consider a version of the model with limited pledgeability of physical assets and heterogeneous idiosyncratic volatility.

Limited pleadgeability of capital. Let $b_{i,t} \equiv n_{i,t} - q_t k_t$ denote the amount of safe assets held by entrepreneur i. The natural borrowing limit can be written as

$$-b_{i,t} \leq h_{i,t} + qk_{i,t}. \tag{D.1}$$

Entrepreneur can borrow freely against physical assets or human wealth. Let's now assume that there is limited pledgeability of physical assets, that is, entrepreneurs can only borrow a fraction of $1 - \lambda^{-1}$ of the value of physical assets, a form of collateral constraint:

$$-b_{i,t} \geq h_{i,t} + (1 - \lambda^{-1})qk_{i,t} \Rightarrow qk_{i,t} \leq \lambda\omega_{i,t}, \tag{D.2}$$

where $\lambda \geq 1$. Hence, the entrepreneur faces a portfolio problem subject to leverage constraints. The HJB for an entrepreneur can be written as[33]

$$\begin{aligned}\frac{\rho}{1-\gamma} = \max_{c_{i,t},k_{i,t},l_{i,t},\iota_{i,t},\theta^{ag}_{i,t},\theta^{id}_{i,t}} \Bigg\{ & \frac{\zeta^\gamma(a)}{1-\gamma}\left(\frac{c_{i,t}}{\omega_{i,t}}\right)^{1-\gamma} + r + \frac{q_t k_{i,t}}{\omega_{i,t}}(\mu^R_{i,t} - r_t) - \frac{p^{ag}\theta^{ag}_{i,t}}{\omega_{i,t}} + \frac{h_{i,t}}{\omega_{i,t}}\sigma_A p^{ag} - \frac{c_{i,t}}{\omega_{i,t}} \\ & - \frac{\gamma}{1-\gamma}\frac{1}{\zeta(a)}\frac{\partial \zeta(a)}{\partial a} - \frac{\gamma}{2}\left[\left(\frac{q_t k_{i,t} + h_{i,t}}{\omega_{i,t}}\sigma_A - \frac{\theta^{ag}_{i,t}}{\omega_{i,t}}\right)^2 + \left(\frac{q_t k_{i,t}}{\omega_{i,t}}\sigma_{I,i} - \frac{\theta^{id}_{i,t}}{\omega_{i,t}}\right)^2\right]\Bigg\},\end{aligned}$$

subject to leverage and insurance constraints: $q_t k_{i,t} \leq \lambda\omega_{i,t}$ and $\theta^{id}_{i,t} \leq (1-\phi)q_t k_{i,t}\sigma_{I,i}$.

The optimal capital demand is given by

$$\frac{qk_{i,t}}{\omega_{i,t}} = \frac{\varphi^{id}_i}{\gamma\phi\sigma_{I,i}}, \tag{D.3}$$

where $\varphi^{id}_{i,t}$ satisfies the condition:

$$\varphi^{id}_i = \min\left\{\frac{\mu^R_i - r - p^{ag}\sigma_A}{\phi\sigma_{I,i}}, \gamma\phi\sigma_{I,i}\lambda\right\}.$$

φ^{id}_i corresponds to the Lagrange multiplier on the insurance constraint. In this case, this multiplier may not coincide with the price of idiosyncratic risk $p^{id}_i \equiv \frac{\mu^R_i - r - p^{ag}\sigma_A}{\phi\sigma_{I,i}}$.

Equilibrium implications. We only sketch the equilibrium implications. Let $\overline{k}_j$, $\overline{\omega}_j$, and $\overline{\mu}^R_j$ denote the long-run values of capital, total wealth, and expected marginal return for an entrepreneur of type j. We have seen in Section 5 that expected returns are decreasing

[33] A similar result on the optimal portfolio share with leverage constraints can be found for investors without labor income in, for instance, Grossman and Vila (1992) and Detemple and Murthy (1997).

in the capital stock with DRS, so we can write $\overline{\mu}_j^R = f(\overline{k}_j)$, where $f'(\cdot) < 0$. By analogy with the result in Proposition 2, we can assume that $\overline{\omega}_j$ is an increasing function of the return $\overline{\mu}_j^R$, $\omega_j = g(\overline{\mu}_j^R)$, where $g'(\cdot) > 0$. For a constrained investor, we have $q\overline{k}_j = \lambda\overline{\omega}_j$, so $\overline{k}_j = \frac{\lambda}{q} g_j(\overline{\mu}_j^R)$ The capital stock would be determined by the intersection of $f(\overline{k}_j)$ and $g_j^{-1}(q\overline{k}_j/\lambda)$. In this case, variations in wealth, associated with shifts of the curve $g_j(\cdot)$, will lead to variations in the capital stock and ultimately expected returns. This argument indicates that cross-sectional differences in expected returns, controlling for differences in risk, are driven by differences in net worth for constrained entrepreneurs.

D.2 Endogenous occupational choice

In this subsection, we introduce an occupational choice into the households' problem. Moreover, we assume that wage earners have a finite horizon and imperfect altruism in the same way as entrepreneurs.

D.2.1 The occupational choice

At the beginning of life, a household can choose to become an entrepreneur or a wage earner. To become an entrepreneur, household i must pay a fixed cost $\varphi_i \tilde{y}_{i,t}$, where φ_i is a cost parameter draw from a distribution $F_\varphi(\cdot)$ with support $[\underline{\varphi}, \overline{\varphi}]$. Let $\tilde{V}_t(\tilde{n}_i, a)$ denote the value function of a household that chose to become an entrepreneur and $\tilde{V}_t^w(\tilde{n}_i, a)$ the value function of a household who chose to become a wage earner. In contrast to the model from Section 3, a wage earner lives for T periods and derives the same utility of bequests as entrepreneurs.

A household that inherits financial wealth $\tilde{n}_i$ will choose to become an entrepreneur if

$$\tilde{V}_t(\tilde{n}_i - \varphi_i \tilde{y}_t, 0) > \tilde{V}_t^w(\tilde{n}_i, 0). \tag{D.4}$$

The value function of an entrepreneur can be written, after normalization, as $V(n,a) = \zeta(a)^{-\gamma} \frac{(n+h(a))^{1-\gamma}}{1-\gamma}$. Similarly, the value function of a wage earner can be written as $V^w(n,a) = \zeta_w(a)^{-\gamma} \frac{(n+h_w(a))^{1-\gamma}}{1-\gamma}$.

The condition for becoming an entrepreneur can then be written as

$$\zeta(0)^{\frac{\gamma}{\gamma-1}} (n_i + h(0) - \varphi_i y) > \zeta_w(0)^{\frac{\gamma}{\gamma-1}} (n_i + h_w(0)). \tag{D.5}$$

Rearranging the expression above, we obtain that a household becomes an entrepreneur if $\varphi_i < \varphi^*(n_i)$, where the threshold $\varphi^*(n_i)$ is given by

$$\varphi^*(n_i) \equiv \frac{1}{y}\left(\frac{\zeta(0)}{\zeta_w(0)}\right)^{-\frac{\gamma}{\gamma-1}}\left[\left(\left(\frac{\zeta(0)}{\zeta_w(0)}\right)^{\frac{\gamma}{\gamma-1}} - 1\right) n_i + \left(\frac{\zeta(0)}{\zeta_w(0)}\right)^{\frac{\gamma}{\gamma-1}} h(0) - h_w(0)\right]. \quad \text{(D.6)}$$

It can be shown that $\zeta(0) > \zeta_\omega(0)$ for $\gamma > 1$, so households who received larger bequests are more likely to become entrepreneurs. The difference between $\zeta(0)$ and $\zeta_\omega(0)$ is increasing in p^{id}, the shadow price of idiosyncratic risk.

As the cost parameter is drawn independently of the bequest a household receives, then the mass of entrepreneurs in a stationary equilibrium is given by

$$\chi_e = F_\varphi\left(\varphi^*(n(0))\right), \quad \text{(D.7)}$$

where $n(0)$ is the average financial wealth of newborn entrepreneurs.

In a stationary equilibrium, the mass of entrepreneurs is constant. As $\zeta(0)$, $\zeta_w(0)$, and $n(0)$ depend on the interest rate and the aggregate and idiosyncratic risk premia, then the share of entrepreneurs in the economy depends on the equilibrium expected returns.

D.2.2 Wage earners' problem and equilibrium determination

The optimal consumption-wealth ratio and demand for insurance for wage earners are now given by

$$\frac{c_{j,t}}{\omega_{j,t}} = \zeta_w(a) = \frac{\bar{r}_w}{1 - \psi e^{-\bar{r}_w(T-a_j)}}, \qquad \theta^{ag}_{j,t} = h_{j,t}\sigma_A - \frac{p^{ag}}{\gamma}\omega_{j,t}, \quad \text{(D.8)}$$

where

$$\bar{r}_w = \frac{1}{\gamma}\rho_w + \left(1 - \frac{1}{\gamma}\right)\left(r + \frac{(p^{ag})^2}{2\gamma}\right). \quad \text{(D.9)}$$

The price of aggregate insurance, wages, and the relative price of capital are the same as in the baseline model:

$$p^{ag} = \gamma\sigma_A, \qquad w = (1-\alpha)K^\alpha, \qquad q = \Phi_0 + \Phi_1(g+\delta). \quad \text{(D.10)}$$

Finite lives for wage earners change the determination of the interest rate. The interest rate is now jointly determined with the capital-labor ratio and the price of idiosyncratic risk by conditions (34), (35), and the market clearing condition for consumption

$$\int_0^T \frac{\bar{r}\omega(a)}{1-\psi e^{-\bar{r}(T-a)}} f(a)da + \int_0^T \frac{\bar{r}_w \omega_w(a)}{1-\psi e^{-\bar{r}_w(T-a)}} f(a)da = \alpha K^\alpha - \iota K, \tag{D.11}$$

where

$$\omega(a) = \omega(0)e^{\left(r+\gamma\sigma_A^2+\frac{(p^{id})^2}{\gamma}-\mu_A-\bar{r}\right)a}\frac{1-\psi e^{-\bar{r}(T-a)}}{1-\psi e^{-\bar{r}T}}, \qquad \omega_w(a) = \omega_w(0)e^{\left(r+\gamma\sigma_A^2-\mu_A-\bar{r}_w\right)a}\frac{1-\psi e^{-\bar{r}_w(T-a)}}{1-\psi e^{-\bar{r}_w T}}.$$

Assuming finite lives for wage earners would change the calibration of ρ_w but otherwise would not affect our main results.

D.3 Uninsurable labor income risk and borrowing constraints

In this subsection, we introduce uninsurable labor income risk into the entrepreneur's problem. This will enable us to study the implications of both insurance and borrowing constraints on entrepreneurial behavior. In particular, we focus on how the inability to borrow against future income affects the entrepreneur's risk-taking decision.

D.3.1 The entrepreneurs' problem with labor income risk

Entrepreneurs receive labor income $\bar{l}_{i,t}$. With Poisson intensity λ_d, entrepreneurs suffer a "disability" shock that reduces their labor income by a factor $1-\xi_d$, that is, labor income is given by $\bar{l}_{i,t} = \bar{l}(a_i)$ in the no-disability state and $\bar{l}_{i,t} = (1-\xi_d)\bar{l}(a_i)$ in the disability state. The disability shock happens only once in an entrepreneur's lifetime, and it is permanent. When either $\lambda_d = 0$ or $\xi_d = 0$, we recover the model with no labor income risk.

We assume that households are subject to a natural borrowing limit, given by $\tilde{n}_{i,t} \geq -(1-\xi_d)\tilde{h}_{i,t}$.[34] As discussed in Aiyagari (1994), the natural borrowing limit under incomplete markets corresponds to the worst-case scenario of the realization of idiosyncratic shocks, which in our setting corresponds to the disability shock happening immediately. Therefore, under the natural borrowing limit, entrepreneurs can borrow at most a fraction $1-\xi_d$ of the human wealth $\tilde{h}_{i,t}$. Note that, even for an arbitrarily small value of λ_d, the borrowing limit is tighter in the presence of uninsurable income risk. In this case, the parameter ξ_d effectively controls the pledgeability of human wealth.

[34] Kaplan and Violante (2010) argue that the standard incomplete markets model with a natural borrowing limit better captures the degree of partial insurance observed in the data than versions of the model with tighter borrowing limits.

The entrepreneur's problem in the no-disability state is given by

$$\tilde{V}_t(\tilde{n}_i, a_i) = \max_{\tilde{c}_i, \tilde{\theta}^{ag}_i, \tilde{\theta}^{id}_t, k_i, l_i, \iota_i} \mathbb{E}_t \left[\int_t^{t+T_d} e^{-\rho(z-t)} \frac{\tilde{c}_{i,z}^{1-\gamma}}{1-\gamma} dz + e^{-\rho T_d} \tilde{V}^d_{t+T_d}(\tilde{n}_{i,t+T_d}, a_i + T_d) \right], \tag{D.12}$$

subject to non-negativity constraints $c_{i,t}, k_{i,t} \geq 0$, the law of motion of $\tilde{n}_{i,t}$

$$d\tilde{n}_{i,t} = \left[\tilde{n}_{i,t} r + \tilde{q}_t k_{i,t}(\mu^R_{i,t} - r) - p^{ag}\tilde{\theta}^{ag}_{i,t} + \tilde{w}_t \bar{l}_{i,t} - \tilde{c}_{i,t} \right] dt + \left(\tilde{q}_t k_{i,t} \sigma_A - \tilde{\theta}^{ag}_{i,t} \right) dZ_t + \left(\tilde{q}_t k_{i,t} \sigma_I - \tilde{\theta}^{id}_{i,t} \right) dZ_{i,t},$$

and insurance and borrowing constraints

$$\tilde{\theta}^{id}_{i,t} \leq (1-\phi)\tilde{q}_t k_{i,t} \sigma_I, \qquad \tilde{n}_{i,t} \geq -(1-\xi_d)\tilde{h}_{i,t}, \tag{D.13}$$

where $\tilde{h}_{i,t} \equiv \mathbb{E}_t \left[\int_t^{t+T-a_i} \frac{\pi_z}{\pi_t} \tilde{w}_z \bar{l}_i dz \right]$, T_d is the minimum of the (random) arrival time of a Poisson process with intensity $\lambda_d > 0$ and the entrepreneur's life horizon $T - a_i$. If $T_d < T - a_i$, then $\tilde{V}^d(\tilde{n}_i, a_i)$ corresponds to the value function in the disability state. As there is no labor income risk in the disability state, this is equal to the value function derived in Section 3. The value function evaluated at $a_i = T$, $\tilde{V}^d_t(\tilde{n}_i, T)$, is given by the bequest function.

D.3.2 The stationary problem

It is convenient to adopt a change of variables and to detrend the problem. First, we follow Aiyagari (1994) and define total (pledgeable) wealth $\tilde{\omega}_{i,t} = \tilde{n}_{i,t} + (1-\xi_d)\tilde{h}_{i,t}$. This corresponds to the total amount of funds available to an entrepreneur, that is, the sum of the financial wealth and the borrowing limit. In the special case where $\xi_d = 0$, we recover the definition of total wealth given in Section 3. With a slight abuse of notation, we denote the entrepreneur's value function as $\tilde{V}_t(\tilde{\omega}_{i,t}, a_{i,t}) = \tilde{V}(\tilde{\omega}_{i,t}, a_{i,t}; A_t)$. Therefore, the entrepreneur's problem depends on the level of aggregate productivity. The next lemma shows that, despite the presence of uninsurable labor income risk and aggregate shocks, it is possible to detrend the problem and to define a stationary equilibrium in a way analogous to Lemma 1.

Lemma 4. *Suppose the economy is in a stationary equilibrium. Then,*

i. Scaled variables are independent of aggregate productivity, that is, the scaled value function $V(\omega_t, a_t) = \frac{\tilde{V}_t(\omega_t A_t, a_t)}{A_t^{1-\gamma}}$ and the scaled policy functions $c_{i,t} = \frac{\tilde{c}_{i,t}}{A_t}$, $\theta^{id}_{i,t} = \frac{\tilde{\theta}^{id}_{i,t}}{A_t}$, and $\theta^{ag}_{i,t} = \frac{\tilde{\theta}^{ag}_{i,t}}{A_t}$ do not depend on A_t.

ii. *The optimal value of $l_{i,t}$ and $\iota_{i,t}$ are given by (15) and (16), respectively. The insurance constraint is binding, and the shadow price of idiosyncratic insurance is given by (22). The price of aggregate insurance is given by $p^{ag} = \gamma\sigma_A$ and the demand for aggregate insurance is given by $\theta^{ag}_{i,t} = (qk_{i,t} - n_{i,t})\,\sigma_A$.*

iii. *The scaled value function satisfies the HJB equation*

$$\hat{\rho}V = \max_{c_i, \sigma^{id}_{i,t}} \frac{c_t^{1-\gamma}}{1-\gamma} + V_a + V_\omega \left[\hat{r}\omega_{i,t} + p^{id}\sigma^{id}_{i,t} + \xi_d w_t \bar{l}_{i,t} - c_{i,t}\right] + \frac{1}{2}V_{\omega\omega}(\sigma^{id}_{i,t})^2 + \lambda_d \left(V^d - V\right), \tag{D.14}$$

where $\sigma^{id}_{i,t} = qk_{i,t}\phi\sigma_I$, $\hat{\rho} \equiv \rho - (1-\gamma)\left(\mu_A - \frac{\gamma\sigma_A^2}{2}\right)$, and $\hat{r} \equiv r + \gamma\sigma_A^2 - \mu_A$.

Proof. See Appendix D.3.4. □

In the first part of Lemma 4, we show that scaled variables are independent of aggregate productivity. This implies that a stationary distribution of scaled wealth exists and there is no need to approximate the aggregate wealth distribution by a finite number of moments as in Krusell and Smith (1998), despite the presence of aggregate risk and incomplete markets. Two assumptions are important to obtain this result: i) a borrowing limit that is proportional to aggregate income; ii) the probability of switching to the disability state is independent of aggregate shocks. Note that the natural borrowing limit in our setting satisfies the first condition. This result echoes the one obtained by Krueger and Lustig (2010), who derive conditions under which uninsurable labor income risk has no effect on the price of aggregate risk. Consistent with their findings, we show that the price of aggregate insurance is $p^{ag} = \gamma\sigma_A$, the same we would obtain in a complete-markets economy.

The second part of the lemma shows that the solution to the entrepreneur's problem shares many of the features we derived without idiosyncratic labor income risk. Labor demand and investment rates are chosen to maximize expected returns, insurance constraints are always binding, and entrepreneurs with low financial wealth have a positive demand for aggregate insurance.

The third part of Lemma 4 shows how the entrepreneur's problem ultimately reduces to a choice of (scaled) consumption and the exposure to idiosyncratic risk, or equivalently, the amount of capital to be employed in the business. In contrast to the problem in Section 3, a closed-form solution to this problem is not available, as the consumption function is not a linear function of total wealth anymore.

D.3.3 The approximate solution

We consider next an approximate solution to problem (D.14). Despite the well-known lack of closed-form solutions, we can provide an analytical characterization using perturbation techniques. In particular, we extend the methods used in Viceira (2001) to a general equilibrium life-cycle model. This approximate solution will allow us to show analytically how borrowing constraints affect entrepreneurs' risk-taking decisions, which is challenging to obtain from problem (D.14) otherwise.

Let $\omega(a)$ denote the average total wealth of entrepreneurs conditional on age a. We consider a log-linear approximation of the consumption and capital functions around the point $(\overline{\omega}, \overline{a})$:

$$\log c(\omega_{i,t}, a_{i,t}) = \log c(\overline{\omega}, \overline{a}) + \psi_{c,\omega}\hat{\omega}_i + \psi_{c,a}\hat{a}_i + \mathcal{O}\left(\hat{\omega}_i^2, \hat{a}_i^2\right) \tag{D.15}$$

$$\log k(\omega_{i,t}, a_{i,t}) = \log k(\overline{\omega}, \overline{a}) + \psi_{k,\omega}\hat{\omega}_i + \psi_{k,a}\hat{a}_i + \mathcal{O}\left(\hat{\omega}_i^2, \hat{a}_i^2\right) \tag{D.16}$$

where $\hat{\omega}_i \equiv \log \omega_i - \log \overline{\omega}$ and $\overline{\omega} \equiv \exp \mathbb{E}[\log \omega_{i,t} | a_{i,t} = \overline{a}]$, given $0 < \overline{a} < T$.

We can write the expressions above in a more compact form:

$$\hat{c}_{i,t} = \psi_{c,\omega}\hat{\omega}_{i,t} + \psi_{c,a}\hat{a}_{i,t}, \qquad \hat{k}_{i,t} = \psi_{k,\omega}\hat{\omega}_{i,t} + \psi_{k,a}\hat{a}_{i,t}, \tag{D.17}$$

up to first order in $(\hat{\omega}_{i,t}, \hat{a}_{i,t})$, where $\hat{z}_{i,t} \equiv \log z(\omega_{i,t}, a_{i,t}) - \log z(\overline{\omega}, \overline{a})$ for $z \in \{c, k\}$.

Note that this log-linear approximation is only feasible when total wealth is always positive, that is, $\omega_{i,t} > 0$ for all i and t. This condition holds in the solution to the entrepreneurs' problem given the combination of the natural borrowing limit with an unbounded marginal utility as $c_{i,t}$ approaches zero (see, e.g., the discussion in Chamberlain and Wilson 2000). If an entrepreneur were to borrow the maximum amount, such that $\omega_{i,t} = 0$, then consumption would be zero in the advent of a disability shock, a state with infinite marginal utility. An entrepreneur is better off by reducing borrowing and avoiding this possibility.

The next proposition provides a characterization of the first-order approximation of the entrepreneurs' problem.

Proposition 7. *Suppose the economy is in a stationary equilibrium and consider a first-order approximation of the policy functions and wealth dynamics around $(\overline{\omega}, \overline{a})$. Then,*

i. Consumption is an age-dependent concave function of total wealth given by

$$c_{i,t} = f_c(a_{i,t})\omega_{i,t}^{\psi_{c,\omega}}, \tag{D.18}$$

where $0 < \psi_{c,\omega} < 1$ *and* $f_c'(a_{i,t}) > 0$ *for* ξ_d *sufficiently small.*

ii. *The demand for capital is given by*

$$\frac{qk_{i,t}}{n_{i,t}} = \frac{1 + (1-\xi_d)\frac{h_{i,t}}{n_{i,t}}}{\gamma\psi_{c,\omega}} \frac{p^{id}}{\phi\sigma_I}. \tag{D.19}$$

iii. *Log total wealth evolves according to*

$$d\log\omega_{i,t} = [\tilde{\psi}_{\omega,0} + \psi_{\omega,a}a_{i,t} + \psi_{\omega,\omega}\log\omega_{i,t}]\,dt + \frac{p^{id}}{\gamma\psi_{c,\omega}}dZ_{i,t}, \tag{D.20}$$

where $\psi_{\omega,\omega} < 0$ *and* $\psi_{\omega,a} < 0$ *for* ξ_d *sufficiently small.*

Proof. See Appendix D.3.5. □

The first part of Proposition 7 describes the consumption function. Consumption is a strictly concave function of total wealth, as it is typical of problems with uninsurable labor income risk, where $\psi_{c,\omega} \in (0,1)$ represents the elasticity of consumption with respect to total wealth. This is in contrast to the consumption function derived in Proposition 1, where consumption is a linear function of $\omega_{i,t}$, that is, $\psi_{c,\omega} = 1$. For ξ_d sufficiently small, consumption is an increasing function of age given $\omega_{i,t}$, reflecting the impact of entrepreneurs' finite horizon.[35]

The second part of Proposition 7 gives the demand for capital and it corresponds to the main result of this section. The fact that entrepreneurs cannot borrow against future labor income has two (opposite) effects on the demand for capital. First, entrepreneurs have fewer resources available to them when $\xi_d > 0$, which tends to reduce the business scale. This effect is particularly more pronounced for entrepreneurs who are close to the borrowing limit, that is, $-n_{i,t}$ is close to $(1-\xi_d)h_{i,t}$. Second, the limited pledgeability of human wealth *reduces* the effective risk aversion of entrepreneurs, as $\gamma\psi_{c,\omega} < \gamma$. The fraction of human wealth that cannot be used to fund a business investment acts as a buffer against future shocks, which makes the entrepreneur less concerned about taking investment risks.

It can be shown that uninsurable labor income risk reduces the scale of the business

[35]The condition on ξ_d is necessary, as with uninsurable labor income risk the slope of the labor income profile also plays a role. The effect of a finite horizon on consumption is attenuated when labor income declines with age, and this effect is amplified when labor income increases with age.

relative to an economy with $\xi_d = 0$ if and only if

$$n_{i,t} < \left[\frac{\xi_d}{1-\psi_{c,\omega}} - 1\right] h_{i,t}. \tag{D.21}$$

Tighter borrowing constraints, captured by $\xi_d > 0$, reduce investment in the business for poor entrepreneurs, but they increase investment in the business for rich entrepreneurs. Given that these two forces move in opposite directions, the aggregate effect of borrowing constraints tends to be muted. Importantly, for all values of ξ_d, a declining human-financial wealth ratio over the life cycle causes entrepreneurs' exposure to the business to decline with age, as in the baseline model.

Entrepreneurs have an age-dependent target for wealth: $\frac{\tilde{\psi}_{\omega,0}+\psi_{\omega,a}a}{|\psi_{\omega,\omega}|}$. Entrepreneurs build up wealth when $\omega_{i,t}$ is below target and they decumulate wealth when $\omega_{i,t}$ is above target. When $\psi_{\omega,a} < 0$, the target on total wealth drifts down with age, again an implication of the entrepreneurs' finite horizon.

Note that the ratio of consumption to financial wealth can be written as

$$\frac{c_{i,t}}{n_{i,t}} = f_c(a_{i,t})\omega_{i,t}^{-(1-\psi_{c,\omega})}\left(1 + (1-\xi_d)\frac{h_{i,t}}{n_{i,t}}\right). \tag{D.22}$$

As typically $f_c(\cdot)$ is increasing with age and $\omega_{i,t}$ is decreasing with age on average, we obtain that the first two terms in the expression above increase with age. Figure 5 shows that the human-financial wealth ratio is declining with age. In line with our discussion in Section 3, the consumption-financial wealth ratio then depends on two forces that move in opposite directions with age.

Therefore, we conclude that introducing uninsurable labor income risk and borrowing constraints does not change substantially our results, while it adds a significant layer of complexity to the analysis.

D.3.4 Proof of Lemma 4

Proof. With a slight abuse of notation, we denote the entrepreneur's value function as a function of total wealth, age, and aggregate productivity: $\tilde{V}_t(\tilde{\omega}_i, a_i) = \tilde{V}(\tilde{\omega}_i, a_i; A_t)$. The HJB for the entrepreneur's problem is given by

$$\rho\tilde{V}_t = \max_{\tilde{c}_i,\tilde{\theta}_i^{ag},\tilde{\theta}_t^{id},k_i,l_i,\iota_i} \frac{\tilde{c}_t^{1-\gamma}}{1-\gamma} + \frac{\mathbb{E}_t\left[d\tilde{V}_t\right]}{dt} \tag{D.23}$$

subject to the insurance and borrowing constraints (D.13).

We guess and verify that the value function can be written as

$$\tilde{V}(\tilde{\omega}, a; A_t) = A_t^{1-\gamma} V\left(\frac{\tilde{\omega}}{A_t}, a\right), \qquad \tilde{V}^d(\tilde{\omega}, a; A_t) = A_t^{1-\gamma} V^d\left(\frac{\tilde{\omega}}{A_t}, a\right), \tag{D.24}$$

where $V(\omega, a)$ and $V^d(\omega, a)$ are independent of A_t.

Let $\omega_{i,t} \equiv \tilde{\omega}_{i,t}/A_t$ denote scaled total wealth. Using an argument analogous to the one used in Lemma 1, we can derive the law of motion $h_{i,t} = \tilde{h}_{i,t}/A_t$, which then gives the law of motion of $\omega_{i,t}$:

$$d\omega_{i,t} = \mu_{\omega_i,t}dt + \sigma_{i,t}^{ag}dZ_t + \sigma_{i,t}^{id}dZ_{i,t}, \tag{D.25}$$

where $\sigma_{i,t}^{ag} \equiv (qk_{i,t} + (1-\xi_d)h_{i,t} - \omega_{i,t})\,\sigma_A - \theta_{i,t}^{ag}$, $\sigma_{i,t}^{id} \equiv qk_{i,t}\sigma_I - \theta_{i,t}^{id}$, and

$$\mu_{\omega_i,t} \equiv (r + p^{ag}\sigma_A - \mu_A)\omega_{i,t} + qk_{i,t}(\mu_{i,t}^R - r - p^{ag}\sigma_A) + (p^{ag} - \sigma_A)\sigma_{i,t}^{ag} + \xi_d w_t \bar{l}_{i,t} - c_{i,t}. \tag{D.26}$$

The HJB equation for the scaled value function can be written as

$$\begin{aligned}\hat{\rho}V = \max_{c_i,\theta_i^{ag},\theta_t^{id},k_i,l_i,\iota_i} \frac{c_t^{1-\gamma}}{1-\gamma} + V_a + V_\omega\left[\hat{r}\omega_{i,t} + qk_{i,t}(\mu_{i,t}^R - r - p^{ag}\sigma_A) + \hat{p}^{ag}\sigma_{i,t}^{ag} + \xi_d w_t \bar{l}_{i,t} - c_{i,t}\right] + \\ + \frac{1}{2}V_{\omega\omega}\left((\sigma_{i,t}^{ag})^2 + (\sigma_{i,t}^{id})^2\right) + \lambda_d\left(V^d - V\right)\end{aligned} \tag{D.27}$$

subject to $\theta_{i,t}^{id} \le (1-\phi)qk_{i,t}\sigma_I$ and $\omega_{i,t} \ge 0$, where

$$\hat{\rho} \equiv \rho - (1-\gamma)\left(\mu_A - \frac{\gamma\sigma_A^2}{2}\right), \qquad \hat{r} \equiv r + p^{ag}\sigma_A - \mu_A, \qquad \hat{p}^{ag} \equiv p^{ag} - \gamma\sigma_A. \tag{D.28}$$

As the HJB equation above is independent of A_t, we conclude that $V(\omega, a)$ is also independent of A_t, confirming our initial guess.

As $l_{i,t}$ and $\iota_{i,t}$ only enter the problem through $\mu_{i,t}^R$, it is optimal to choose them to maximize expected returns. Expected returns are then constant and equalized across entrepreneurs, so we drop the dependence on entrepreneur and time: $\mu_{i,t}^R = \mu^R$. The first-order condition with respect to $\theta_{i,t}^{ag}$ is given by

$$\hat{p}^{ag} = -\frac{V_{\omega\omega}}{V_\omega}\sigma_{i,t}^{ag} \Rightarrow \theta_{i,t}^{ag} = (qk_{i,t} + (1-\xi_d)h_{i,t} - \omega_{i,t})\,\sigma_A + \frac{V_\omega}{V_{\omega\omega}}\hat{p}^{ag}. \tag{D.29}$$

An argument analogous to the one used for the model without labor income risk establishes that $p^{ag} = \gamma\sigma_A$ in stationary equilibrium. This implies that $\hat{p}^{ag} = 0$ and $\sigma_{i,t}^{ag} = 0$, so entrepreneurs choose the same exposure to aggregate risk in equilibrium.

Let $V_\omega p^{id}$ denote the Lagrange multiplier on the insurance constraint. The first-order conditions with respect to $k_{i,t}$ and $\theta^{id}_{i,t}$ are given by

$$\mu^R_{i,t} - r - \gamma\sigma_A^2 + p^{id}(1-\phi)\sigma_I = -\frac{V_{\omega\omega}}{V_\omega}\left[\sigma^{ag}_{i,t}\sigma_A + \sigma^{id}_{i,t}\sigma_I\right], \qquad p^{id} = -\frac{V_{\omega\omega}}{V_\omega}\sigma^{id}_{i,t}. \quad \text{(D.30)}$$

Given that $\sigma^{id}_{i,t} > 0$ by the insurance constraint and given the concavity of the value function, $V_{\omega\omega} < 0$, we have that $p^{id} > 0$. Therefore, the insurance constraint is always binding, that is, $\theta^{id}_{i,t} = (1-\phi)qk_{i,t}\sigma_I$. Rearranging the expressions above, we obtain

$$qk_{i,t} = -\frac{V_\omega}{V_{\omega\omega}}\frac{p^{id}}{\phi\sigma_I}, \qquad p^{id} = \frac{\mu^R - r - p^{ag}\sigma_A}{\phi\sigma_I}. \quad \text{(D.31)}$$

Finally, the optimality condition for consumption is given by

$$c_{i,t}^{-\gamma} = V_\omega \Rightarrow c(\omega_{i,t}, a_{i,t}) = V_\omega^{-\frac{1}{\gamma}}(\omega_{i,t}, a_{i,t}). \quad \text{(D.32)}$$

□

D.3.5 Proof of Proposition 7

Proof. We provide next a complete characterization of the first-order approximation of the entrepreneurs' problem. We proceed in four steps. First, we derive the law of motion of $\hat{\omega}_{i,t}$. Second, we solve for the demand for capital. Third, we will solve for the consumption function. Fourth, we derive the conditions determining the approximation point $\overline{\omega}$.

Law of motion of the state. The log of total wealth for entrepreneur i evolves according to

$$d\log\omega_{i,t} = \left[\hat{r} + \frac{qk_{i,t}}{\omega_{i,t}}p^{id}\phi\sigma_I - \frac{1}{2}\left(\frac{qk_{i,t}}{\omega_{i,t}}\phi\sigma_I\right)^2 + \xi_d\frac{w\bar{l}_{i,t}}{\omega_{i,t}} - \frac{c_{i,t}}{\omega_{i,t}}\right]dt + \frac{qk_{i,t}}{\omega_{i,t}}\phi\sigma_I dZ_{i,t}. \quad \text{(D.33)}$$

Log-linearizing the law of motion of $\omega_{i,t}$, we obtain

$$d\log\omega_{i,t} = \left[\hat{r} + \frac{q\bar{k}}{\bar{\omega}} p^{id}\phi\sigma_I(1+\hat{k}_{i,t}-\hat{\omega}_{i,t}) - \frac{1}{2}\left(\frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\right)^2\left(1+2(\hat{k}_{i,t}-\hat{\omega}_{i,t})\right)\right.$$
$$\left.+\xi_d\frac{w\bar{l}(\bar{a})}{\bar{\omega}}\left(1+\frac{\bar{l}'(\bar{a})}{\bar{l}(\bar{a})}\hat{a}_{i,t}-\hat{\omega}_{i,t}\right) - \frac{\bar{c}}{\bar{\omega}}(1+\hat{c}_{i,t}-\hat{\omega}_{i,t})\right]dt + \frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\left(1+\hat{k}_{i,t}-\hat{\omega}_{i,t}\right)dZ_{i,t}. \tag{D.34}$$

Rearranging the expression above, we get

$$d\hat{\omega}_{i,t} = \left[\left(\frac{q\bar{k}}{\bar{\omega}}p^{id}\phi\sigma_I - \left(\frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\right)^2\right)\left(\hat{k}_{i,t}-\hat{\omega}_{i,t}\right) + \xi_d\frac{w\bar{l}(\bar{a})}{\bar{\omega}}\left(\frac{\bar{l}'(\bar{a})}{\bar{l}(\bar{a})}\hat{a}_{i,t}-\hat{\omega}_{i,t}\right) - \frac{\bar{c}}{\bar{\omega}}(\hat{c}_{i,t}-\hat{\omega}_{i,t})\right]dt$$
$$+\mathbb{E}[d\log\omega_{i,t}|a_{i,t}=\bar{a}] + \frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\left(1+\hat{k}_{i,t}-\hat{\omega}_{i,t}\right)dZ_{i,t}. \tag{D.35}$$

We can write the expression above in more compact form:

$$d\hat{\omega}_{i,t} = [\psi_{\omega,0}+\psi_{\omega,a}\hat{a}_{i,t}+\psi_{\omega,\omega}\hat{\omega}_{i,t}]\,dt + \left[\psi^{\sigma}_{\omega,0}+\psi^{\sigma}_{\omega,a}\hat{a}_{i,t}+\psi^{\sigma}_{\omega,\omega}\hat{\omega}_{i,t}\right]dZ_{i,t}, \tag{D.36}$$

where

$$\psi^{\sigma}_{\omega,0} \equiv \frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I, \qquad \psi^{\sigma}_{\omega,\omega} \equiv \psi^{\sigma}_{\omega,0}(\psi_{k,\omega}-1), \qquad \psi^{\sigma}_{\omega,a} \equiv \psi^{\sigma}_{\omega,0}\psi_{k,a}. \tag{D.37}$$

and

$$\psi_{\omega,0} = \hat{r} + \frac{q\bar{k}}{\bar{\omega}}p^{id}\phi\sigma_I - \frac{1}{2}\left(\frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\right)^2 + \xi_d\frac{w\bar{l}(\bar{a})}{\bar{\omega}} - \frac{\bar{c}}{\bar{\omega}} \tag{D.38}$$

$$\psi_{\omega,\omega} = \left(\frac{q\bar{k}}{\bar{\omega}}p^{id}\phi\sigma_I - \left(\frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\right)^2\right)(\psi_{k,\omega}-1) - \xi_d\frac{w\bar{l}(\bar{a})}{\bar{\omega}} - \frac{\bar{c}}{\bar{\omega}}(\psi_{c,\omega}-1), \tag{D.39}$$

$$\psi_{\omega,a} = \left(\frac{q\bar{k}}{\bar{\omega}}p^{id}\phi\sigma_I - \left(\frac{q\bar{k}}{\bar{\omega}}\phi\sigma_I\right)^2\right)\psi_{k,a} + \xi_d\frac{w\bar{l}'(\bar{a})}{\bar{\omega}} - \frac{\bar{c}}{\bar{\omega}}\psi_{c,a}, \tag{D.40}$$

Demand for capital. The first-order condition for capital and consumption are given by

$$-V_{\omega\omega}\sigma^{id}_{i,t} = V_{\omega}p^{id}, \qquad V_{\omega} = c^{-\gamma}_{i,t}. \tag{D.41}$$

Using the fact that $dV_{\omega}dZ_{i,t} = V_{\omega\omega}\sigma^{id}_{i,t}dt$ and the expressions above, we can express the optimality condition for capital as follows:

$$p^{id}dt = \gamma \frac{dc_{i,t}}{c_{i,t}} dZ_{i,t}. \tag{D.42}$$

Up to first order, we have that $\frac{dc_{i,t}}{c_{i,t}} dZ_{i,t} = d\hat{c}_{i,t} dZ_{i,t} = \psi_{c,\omega} d\hat{\omega}_{i,t} dZ_{i,t}$. We can then write the expression above as follows:

$$p^{id} = \gamma \psi_{c,\omega} \left[\psi^{\sigma}_{\omega,0} + \psi^{\sigma}_{\omega,a} \hat{a}_{i,t} + \psi^{\sigma}_{\omega,\omega} \hat{\omega}_{i,t} \right]. \tag{D.43}$$

As the expression above must hold for all $\hat{a}$ and $\hat{\omega}_{i,t}$, then we must have $\psi^{\sigma}_{\omega,\omega} = \psi^{\sigma}_{\omega,a} = 0$. This implies that coefficients in the expansion for capital are given by

$$\psi_{k,\omega} = 1, \qquad \psi_{k,a} = 0. \tag{D.44}$$

The exposure to the business relative to total wealth is then the same for all entrepreneurs, in line with the results in Section 3:

$$\frac{qk_{i,t}}{\omega_{i,t}} = \frac{q\overline{k}}{\overline{\omega}}. \tag{D.45}$$

Using the fact that $V_{\omega\omega} = -\gamma c^{-\gamma-1} c_{\omega}$ and evaluating the first-order condition for capital at $(\overline{\omega}, \overline{a})$, we obtain

$$\gamma \frac{c_{\omega}(\overline{\omega}, \overline{a})}{c(\overline{\omega}, \overline{a})} qk(\overline{\omega}, \overline{a}) \phi \sigma_I = p^{id} \Rightarrow \frac{qk(\overline{\omega}, \overline{a})}{\overline{\omega}} = \frac{1}{\gamma \psi_{c,\omega}} \frac{p^{id}}{\phi \sigma_I}, \tag{D.46}$$

where we used that $\psi_{c,\omega} = \frac{c_{\omega}(\overline{\omega}, \overline{a})\overline{\omega}}{c(\overline{\omega}, \overline{a})}$.

The demand for capital can then be written as

$$\frac{qk_{i,t}}{n_{i,t}} = \frac{1 + (1 - \xi_d)\frac{h_{i,t}}{n_{i,t}}}{\gamma \psi_{c,\omega}} \frac{p^{id}}{\phi \sigma_I}. \tag{D.47}$$

Consumption. The envelope condition with respect to ω for problem (D.14) is given by

$$\hat{\rho} V_{\omega} = V_{\omega} \hat{r} + \frac{\mathbb{E}\left[dV_{\omega}\right]}{dt}. \tag{D.48}$$

Using Ito's lemma, we can write the expression above as follows

$$r = \rho + \gamma \left(\mu_A + \frac{1}{dt}\mathbb{E}[d\hat{c}_{i,t}] \right) - \frac{\gamma(\gamma+1)}{2}\sigma_A^2 - \frac{\gamma^2}{2}\mathbb{E}[d\hat{c}_{i,t}^2] - \lambda_d \left(\left(\frac{c_{i,t}^d}{c_{i,t}} \right)^{-\gamma} - 1 \right). \quad \text{(D.49)}$$

Up to first order, the expression above can be written as

$$\gamma\psi_{c,\omega}\left(\psi_{\omega,\omega}\hat{\omega}_{i,t} + \psi_{\omega,a}\hat{a}_{i,t}\right) + \gamma\lambda_d \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} \left(\hat{\omega}_{i,t}(1-\psi_{c,w}) + \left(\frac{\zeta'(\overline{a})}{\zeta(\overline{a})} - \psi_{c,a} \right)\hat{a} \right) = constant. \quad \text{(D.50)}$$

As the expression above must hold for all values of $\hat{\omega}_{i,t}$ and $\hat{a}_{i,t}$, we obtain $\psi_{\omega,\omega} = \psi_{\omega,a} = 0$. This implies the following conditions must hold:

$$\psi_{c,\omega}\psi_{\omega,\omega} + \lambda_d \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} (1-\psi_{c,\omega}) = 0, \qquad \psi_{c,\omega}\psi_{\omega,a} + \lambda_d \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} \left(\frac{\zeta'(\overline{a})}{\zeta(\overline{a})} - \psi_{c,a} \right) = 0. \quad \text{(D.51)}$$

Using the expression for $\psi_{\omega,\omega}$, we obtain a quadratic equation for $\psi_{c,\omega}$:

$$\psi_{c,\omega}^2 - \psi_{c,\omega}\left[1 - \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} \right] - \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} = 0. \quad \text{(D.52)}$$

The equation above has a positive and a negative solution, where the economically relevant solution is the positive one:

$$\psi_{c,\omega} = \frac{1}{2}\left[1 - \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} + \sqrt{ \left(1 - \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} \right)^2 + 4\lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} } \right]. \quad \text{(D.53)}$$

We will show next that $\psi_{c,\omega} < 1$. Assuming that $\psi_{c,\omega} > 1$, the expression above implies that

$$\sqrt{ \left(1 - \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} \right)^2 + 4\lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} } > 1 + \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} + \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}}. \quad \text{(D.54)}$$

Squaring both sides of the inequality above, we obtain

$$\left(1 - \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} \right)^2 + 4\lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} > \left(1 + \lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} + \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} \right)^2. \quad \text{(D.55)}$$

Rearranging the expression above, we obtain

$$4\lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} > 4 \left(\lambda_d \frac{\overline{\omega}}{\overline{c}} \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} + \xi_d \frac{w\bar{l}(\overline{a})}{\overline{c}} \right), \tag{D.56}$$

which is a contradiction. Therefore, we must have $0 < \psi_{c,\omega} < 1$.

We solve next for $\psi_{c,a}$. The coefficient $\psi_{c,a}$ satisfies the equation:

$$\psi_{c,\omega} \left(\xi_d \frac{w\bar{l}'(\overline{a})}{\overline{\omega}} - \frac{\overline{c}}{\overline{\omega}} \psi_{c,a} \right) + \lambda_d \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} \left(\frac{\zeta'(\overline{a})}{\zeta(\overline{a})} - \psi_{c,a} \right) = 0. \tag{D.57}$$

Rearranging the expression above, we obtain

$$\psi_{c,a} = \frac{\psi_{c,\omega} \frac{\overline{c}}{\overline{\omega}} \xi_d \frac{w\bar{l}'(\overline{a})}{\overline{c}} + \lambda_d \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} \frac{\zeta'(\overline{a})}{\zeta(\overline{a})}}{\psi_{c,\omega} \frac{\overline{c}}{\overline{\omega}} + \lambda_d \left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma}}. \tag{D.58}$$

For ξ_d sufficiently small, the expression above is positive, as $\zeta'(a)$, the consumption-wealth ratio in the absence of labor income risk, is positive.

We can then write consumption as follows

$$c_{i,t} = c(\overline{\omega}, \overline{a}) e^{\psi_{c,a}(a_{i,t} - \overline{a})} \left(\frac{\omega_{i,t}}{\overline{\omega}} \right)^{\psi_{c,\omega}} \equiv f_c(a_{i,t}) \omega_{i,t}^{\psi_{c,\omega}}, \tag{D.59}$$

where the second equality defines the age-dependent function $f_c(a)$. Note that $f_c(a)$ is increasing in a for ξ_d sufficiently small.

It remains to solve for the consumption-wealth ratio at $(\overline{\omega}, \overline{a})$. From the envelope condition, we obtain

$$r = \rho + \gamma \left(\mu_A + \psi_{c,\omega}\psi_{\omega,0} + \psi_{c,a} \right) - \frac{\gamma(\gamma+1)}{2}\sigma_A^2 - \frac{\gamma^2}{2}(\psi_{c,\omega}\psi_{\omega,0}^{\sigma})^2 - \lambda_d \left(\left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - 1 \right). \tag{D.60}$$

where $\psi_{\omega,0} = r + \gamma\sigma_A^2 - \mu_A + \frac{q\bar{k}}{\overline{\omega}} p^{id} \phi\sigma_I - \frac{1}{2} \left(\frac{q\bar{k}}{\overline{\omega}} \phi\sigma_I \right)^2 + \xi_d \frac{w\bar{l}(\overline{a})}{\overline{\omega}} - \frac{\overline{c}}{\overline{\omega}}$.

Rearranging the expression above, we obtain

$$\psi_{c,a} = \psi_{c,\omega} \left(\frac{\overline{c}}{\overline{\omega}} - \xi_d \frac{w\bar{l}(\overline{a})}{\overline{\omega}} \right) - \overline{r}_d - (1 - \psi_{c,\omega}) \left(\mu_A - \frac{\gamma\sigma_A^2}{2} \right) + \frac{\lambda_d}{\gamma} \left(\left(\frac{\zeta(\overline{a})\overline{\omega}}{\overline{c}} \right)^{-\gamma} - 1 \right), \tag{D.61}$$

where

$$\bar{r}_d \equiv \frac{1}{\gamma}\rho + \left(\psi_{c,\omega} - \frac{1}{\gamma}\right)\left(r + \frac{(p^{ag})^2}{2\gamma} + \frac{(p^{id})^2}{2\gamma\psi_{c,\omega}}\right). \tag{D.62}$$

Using the expressions for $\psi_{c,\omega}$ and $\psi_{c,a}$, we obtain a non-linear equation for $\bar{c}$ given $\overline{\omega}$. Note that if $\xi_d = \lambda_d = 0$, we recover a linearized version of the equation determining $\frac{\zeta'(a)}{\zeta(a)}$, Equation (A.18).

Wealth dynamics and the approximation point. The law of motion of $\hat{\omega}_{i,t}$ can be written as

$$d\hat{\omega}_{i,t} = [\psi_{\omega,0} + \psi_{\omega,a}\hat{a}_{i,t} + \psi_{\omega,\omega}\hat{\omega}_{i,t}]\,dt + \frac{p^{id}}{\gamma\psi_{c,\omega}}dZ_{i,t}, \tag{D.63}$$

where

$$\psi_{\omega,\omega} = \frac{\bar{c}}{\overline{\omega}}\left[1 - \psi_{c,\omega} - \xi_d\frac{w\bar{l}(\bar{a})}{\bar{c}}\right], \qquad \psi_{\omega,a} = \lambda_d\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma}\frac{\bar{c}}{\overline{\omega}}\frac{\xi_d\frac{w\bar{l}'(\bar{a})}{\bar{c}} - \frac{\zeta'(\bar{a})}{\zeta(\bar{a})}}{\psi_{c,\omega}\frac{\bar{c}}{\overline{\omega}} + \lambda_d\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma}}. \tag{D.64}$$

We will show next that $\psi_{\omega,\omega} < 0$. For the sake of reaching a contradiction, assume that $\psi_{\omega,\omega} \geq 0$. This implies that the following condition is satisfied:

$$1 + \lambda_d\frac{\overline{\omega}}{\bar{c}}\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma} - \xi_d\frac{w\bar{l}(\bar{a})}{\bar{c}} \geq \sqrt{\left(1 - \lambda_d\frac{\overline{\omega}}{\bar{c}}\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma} - \xi_d\frac{w\bar{l}(\bar{a})}{\bar{c}}\right)^2 + 4\lambda_d\frac{\overline{\omega}}{\bar{c}}\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma}} \tag{D.65}$$

Squaring both sides of the inequality and rearranging, we obtain

$$4\lambda_d\frac{\overline{\omega}}{\bar{c}}\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma}\left(1 - \xi_d\frac{w\bar{l}(\bar{a})}{\bar{c}}\right) \geq 4\lambda_d\frac{\overline{\omega}}{\bar{c}}\left(\frac{\zeta(\bar{a})\overline{\omega}}{\bar{c}}\right)^{-\gamma}, \tag{D.66}$$

which is a contradiction. Thus, we must have $\psi_{\omega,\omega} < 0$.

We solve next for $\overline{\omega}$. Let $\overline{\omega}(a) \equiv \exp\left(\mathbb{E}[\log\omega_{i,t}|a_{i,t} = a]\right)$, then

$$\frac{d\log\overline{\omega}(a)}{da} = \psi_{\omega,0} + \psi_{\omega,a}(a - \bar{a}) + \psi_{\omega,\omega}(\log\overline{\omega}(a) - \log\overline{\omega}). \tag{D.67}$$

Solving the differential equation above, we obtain

$$\log\overline{\omega}(a) = e^{\psi_{\omega,\omega}a}\log\overline{\omega}(0) + \int_0^a\left[\tilde{\psi}_{\omega,0} + \psi_{\omega,a}a'\right]e^{\psi_{\omega,\omega}(a-a')}da'. \tag{D.68}$$

Using the fact that the wealth of entrepreneurs at date T is left as a bequest to the next generation, we can pin down $\overline{\omega}(0)$. Evaluating the expression above at $a = \bar{a}$, for a given reference age $\bar{a}$, we obtain an equation for $\overline{\omega}$. □

D.4 Transitional dynamics with Epstein-Zin preferences

In this extension, we describe the computation of the transitional dynamics with Epstein-Zin preferences. To ensure that that the interest rate is constant, we assume that wage earners have Epstein-Zin preferences and take the limit as the elasticity of intertemporal substitution goes to infinity, that is, workers have linear intertemporal preferences.[36]

Wage earners have the continuous-time analog of Epstein-Zin preferences with EIS ψ_w and risk aversion γ. The wage earner's problem is given by

$$\tilde{V}_t^w(\tilde{n}_j) = \max_{\tilde{c}_j, \tilde{\theta}_j^{ag}} \mathbb{E}_t \left[\int_t^\infty f_w(\tilde{c}_{j,z}, \tilde{V}_z) dz \right], \tag{D.69}$$

subject to $\tilde{n}_{j,t} \geq \tilde{h}_{j,t}$, where $\tilde{h}_{j,t}$ denotes wage earner j's human wealth, non-negativity constraint $\tilde{c}_{j,t} \geq 0$, and the law of motion of financial wealth $\tilde{n}_{j,t}$

$$d\tilde{n}_{j,t} = \left[\tilde{n}_{j,t} r_t - p_t^{ag} \tilde{\theta}_{j,t}^{ag} + \tilde{w}_t \bar{l}_{j,t} - \tilde{c}_{j,t} \right] dt - \tilde{\theta}_{j,t}^{ag} dZ_t,$$

where $f_w(\tilde{c}, V)$ is the aggregator given by

$$f_w(\tilde{c}, \tilde{V}) = \rho_w \frac{(1-\gamma)\tilde{V}}{1-\psi_w^{-1}} \left\{ \left(\frac{\tilde{c}}{((1-\gamma)\tilde{V})^{\frac{1}{1-\gamma}}} \right)^{1-\psi_w^{-1}} - 1 \right\}. \tag{D.70}$$

It is convenient to work with the scaled value function $V_t^w(n)$, which satisfies the condition $\tilde{V}_t^w(\tilde{n}) = A_t^{1-\gamma} V_t^w \left(\frac{\tilde{n}_t}{A_t} \right)$, where $V_t^w(\cdot)$ is independent of A_t. The HJB equation in terms of the scaled value function is given by

$$\begin{aligned} \tilde{\rho}_w \frac{(1-\gamma)V^w}{1-\psi_w^{-1}} = \max_{\tilde{c}, \tilde{\theta}^{ag}} \rho_w \frac{(1-\gamma)V^w}{1-\psi_w^{-1}} \left(\frac{c}{((1-\gamma)V^w)^{\frac{1}{1-\gamma}}} \right)^{1-\psi_w^{-1}} + V_t^w + V_n^w \left[\tilde{r} n + (\gamma\sigma_A - p^{ag})\theta^{ag} + w\bar{l} - c \right] \\ + \frac{1}{2} V_{nn}^w \left(\theta^{ag} + n\sigma_A \right)^2. \end{aligned} \tag{D.71}$$

where $\tilde{\rho}_w \equiv \rho_w - (1-\psi_w^{-1}) \left(\mu_A - \frac{\gamma\sigma_A^2}{2} \right)$ and $\tilde{r}_t \equiv r_t + \gamma\sigma_A^2 - \mu_A$.

[36] This assumption is meant to capture, in an extreme form, the essence of the macro-finance literature that assumes a high EIS (see, e.g., Bansal and Yaron 2004 and Barro 2009). In these models, the high EIS dampens movements in interest rates, so risk premia accounts for most of the variation in discount rates.

Policy functions. The first-order conditions for this problem are given by

$$\rho_w \left((1-\gamma)V^w\right)^{\frac{\psi_w^{-1}-\gamma}{1-\gamma}} c^{-\psi_w^{-1}} = V_n^w, \qquad \gamma\sigma_A - p^{ag} = -\frac{V_{nn}^w}{V_n^w}\left(\theta^{ag} + n\sigma_A\right). \tag{D.72}$$

We will guess and verify that the scaled value function takes the form

$$V_t^w(n_{j,t}) = \left(\frac{\zeta_{w,t}}{\rho_w^{\psi_w}}\right)^{\frac{1-\gamma}{1-\psi_w}} \frac{(n_{j,t}+h_{j,t})^{1-\gamma}}{1-\gamma}, \tag{D.73}$$

where $\zeta_{w,t}$ and $h_{j,t}$ are potentially time-varying, but they are non-stochastic.

Using the expression for the value function above, we obtain the policy functions:

$$\frac{c_{j,t}}{n_{j,t}+h_{j,t}} = \zeta_{w,t}, \qquad \theta_{j,t}^{ag} = \sigma_A h_{j,t} - \frac{p_t^{ag}}{\gamma}(n_{j,t}+h_{j,t}). \tag{D.74}$$

Inserting the policy functions derived above back into the HJB equation, we obtain

$$\frac{\tilde{\rho}_w}{1-\psi_w^{-1}} = \frac{\psi_w^{-1}}{1-\psi_w^{-1}}\zeta_{w,t} + \frac{1}{1-\psi_w}\frac{\dot{\zeta}_{w,t}}{\zeta_{w,t}} + \tilde{r}_t. \tag{D.75}$$

Rearranging the expression above for the case of a stationary equilibrium, so $\dot{\zeta}_{w,t} = 0$, we obtain the consumption-wealth ratio:

$$\zeta_{w,t} = \psi_w\rho_w + (1-\psi_w)\left(r + \frac{\gamma\sigma_A^2}{2}\right), \tag{D.76}$$

which coincides with the expression for the consumption-wealth ratio for wage earners given in (C.6) in the special case of CRRA preferences, i.e., $\psi_w^{-1} = \gamma$.

Price of aggregate risk and interest rate. The demand for aggregate insurance derived above coincides with the expression for aggregate insurance in the CRRA case (see Equation A.23). Therefore, the same argument used in Section C.1 to solve for the price of aggregate risk can be applied in the case of Epstein-Zin preferences in a non-stationary setting. This implies that the price of aggregate risk is constant and given by $p_t^{ag} = \gamma\sigma_A$ during the transitional dynamics.

We consider next the behavior of the interest rate. A derivation analogous to the one in Section C.1 shows that the interest rate in a stationary equilibrium is given by

$$r = \rho_w + \psi_w^{-1} \mu_A - (1 + \psi_w^{-1}) \frac{\gamma \sigma_A^2}{2}, \tag{D.77}$$

which coincides with the expression in Section 6 when $\psi_w^{-1} = \gamma$.

Taking the limit of (D.75) as $\psi_w \to \infty$, we obtain r_t in the case of a non-stationary equilibrium:

$$r_t + \gamma \sigma_A^2 - \mu_A = \rho_w + \frac{\gamma \sigma_A^2}{2} - \mu_A \Rightarrow r_t = \rho_w - \frac{\gamma \sigma_A^2}{2}. \tag{D.78}$$

Therefore, the interest rate is constant when wage earners have linear intertemporal preferences. Moreover, the expression above coincides with the one for the interest rate in the stationary equilibrium (D.77) when specialized to $\psi_w^{-1} = 0$.

E The Optimal Contract

In this appendix, we detail the contracting problem underlying Section 3. We show that the market structure assumed in the main text—where entrepreneurs have access to a riskless bond and both aggregate and idiosyncratic insurance—corresponds to one implementation of the optimal contract allocation. We also show that an alternative environment with a flexible credit line delivers another implementation of the same allocation. The derivation closely follows Di Tella (2017) and is included here for completeness.

E.1 Moral hazard

We assume that aggregate productivity Z_t and the cumulative return $R_{i,t}$ are publicly observable, but the idiosyncratic shock $Z_{i,t}$ is privately observed by entrepreneur i. In addition, the entrepreneur may secretly divert capital at rate $\varsigma_{i,t}$. The return on the business is then

$$dR_{i,t} = \left[\frac{y_{i,t} - w_t l_{i,t} - \iota_{i,t} k_{i,t}}{q_t k_{i,t}} + \frac{\dot{q}_t}{q_t} + \mu_A + \Phi(\iota_{i,t}) - \delta - \varsigma_{i,t} \right] dt + \sigma_A dZ_t + \sigma_I dZ_{i,t}. \tag{E.1}$$

Because $Z_{i,t}$ and $\varsigma_{i,t}$ are privately observed, a principal cannot verify whether low returns stem from a negative shock or from diversion. The optimal contract ensures that $\varsigma_{i,t} = 0$ is incentive-compatible.[37] Expected returns then coincide with condition (9) in the no-stealing case.

[37] This is standard in cash-flow diversion models; see DeMarzo and Sannikov (2006) and DeMarzo and Fishman (2007).

Diverted capital can be sold in the market, but a fraction $1-\phi$ is lost in the process. Proceeds are invested in a hidden account remunerated at the risk-free rate r_t. The entrepreneur's hidden savings $\tilde{S}_{i,t}$ evolve as

$$d\tilde{S}_{i,t} = r_t \tilde{S}_{i,t}\,dt + \phi \tilde{q}_t k_{i,t} \varsigma_{i,t}\,dt. \tag{E.2}$$

E.2 The optimal contract problem

Consumption, investment, and factor demands are contractible. A contract between the principal and entrepreneur is then given by $(\tilde{c}_i, \iota_i, k_i, l_i, \tilde{F}_i)$, where all variables are adapted to the filtration generated by (Z, R_i), and $\tilde{F}_{i,t}$ denotes transfers to the principal. Entrepreneurs cannot commit to long-term contracts. At any point, they can settle outstanding promises, move funds from the hidden account into their bank account, and offer the principal a new contract. Contracts are therefore short-term.

The principal's continuation value. The continuation value to the principal is

$$\tilde{J}_{i,t} = \mathbb{E}_t\left[\int_t^{s_i+T} \frac{\pi_z}{\pi_t}\tilde{F}_{i,z}\,dz\right], \tag{E.3}$$

where the expectation is taken under no stealing ($\varsigma_i = 0$). The principal's SDF π_t evolves according to

$$d\pi_t = -r_t\pi_t\,dt - p_t^{ag}\pi_t\,dZ_t,$$

given the processes for r_t and p_t^{ag}.[38]

To compute the law of motion of $\tilde{J}_{i,t}$, define the martingale

$$G_{i,t} = \int_{s_i}^t \pi_z\,d\tilde{F}_{i,z} + \mathbb{E}_t\left[\int_t^{s_i+T}\pi_z\,d\tilde{F}_{i,z}\right], \tag{E.4}$$

so that $G_{i,s_i} = \mathbb{E}_{s_i}[G_{i,t}]$. By the martingale representation theorem, there exist processes $\sigma^Z_{G_i,t}$ and $\sigma^R_{G_i,t}$ such that

$$\pi_t\tilde{F}_{i,t}\,dt + d(\pi_t\tilde{J}_{i,t}) = \pi_t\sigma^Z_{G_i,t}dZ_t + \pi_t\sigma^R_{G_i,t}\left(dR_{i,t} - \mathbb{E}_t[dR_{i,t}]\right), \tag{E.5}$$

[38] When contracting with a wage earner, the relevant SDF is $\pi_t = e^{-\rho_w t}\tilde{c}_{j,t}^{-\gamma}$. In a stationary equilibrium with $d\tilde{c}_{j,t} = \mu_A\tilde{c}_{j,t}\,dt + \sigma_A\tilde{c}_{j,t}\,dZ_t$, we have $d\pi_t = -\left[\rho_w + \gamma\mu_A - \frac{\gamma(\gamma+1)}{2}\sigma_A^2\right]dt - \gamma\sigma_A dZ_t$, where the drift corresponds to the interest rate r_t and the diffusion to the price of aggregate risk p_t^{ag}.

where $\sigma^Z_{G_i,t}$ represent exposure of $G_{i,t}$ to aggregate shocks and $\sigma^R_{G_i,t}$ represent exposure to return shocks, which include idiosyncratic shocks.

Applying Itô's lemma to $\pi_t \tilde{J}_{i,t}$ and combining with the expression above yields

$$d\tilde{J}_{i,t} = \left[r_t \tilde{J}_{i,t} + p^{ag}_t (\sigma^Z_{J_i,t} + \sigma^R_{J_i,t}\sigma_A) - \tilde{F}_{i,t} \right] dt + \sigma^Z_{J_i,t} dZ_t + \sigma^R_{J_i,t} \left(dR_{i,t} - \mathbb{E}_t[dR_{i,t}] \right), \tag{E.6}$$

where $\sigma^Z_{J_i,t} = \sigma^Z_{G_i,t} + p^{ag}_t \tilde{J}_{i,t}$ and $\sigma^R_{J_i,t} = \sigma^R_{G_i,t}$.

Entrepreneur's financial wealth. The financial wealth of entrepreneur i is defined as

$$\tilde{n}_{i,t} \equiv \tilde{b}_{i,t} + \tilde{q}_t k_{i,t} + \tilde{S}_{i,t} - \tilde{J}_{i,t},$$

i.e., the sum of risk-free asset holdings $\tilde{b}_{i,t}$ the business value $\tilde{q}_t k_{i,t}$, and hidden savings net of promised payments to the principal $\tilde{J}_{i,t}$. Its law of motion is

$$d\tilde{n}_{i,t} = \left[r_t \tilde{b}_{i,t} + \tilde{w}_t \bar{l}_{i,t} - \tilde{F}_{i,t} - \tilde{c}_{i,t} \right] dt + \tilde{q}_t k_{i,t}\, dR_{i,t} + d\tilde{S}_{i,t} - d\tilde{J}_{i,t}. \tag{E.7}$$

Combining (E.6) and (E.7), we obtain

$$\begin{aligned} d\tilde{n}_{i,t} = &\left[r_t(\tilde{n}_{i,t} - \tilde{q}_t k_{i,t}) + \tilde{q}_t k_{i,t} \mu^R_{i,t} - p^{ag}_t (\sigma^Z_{J_i,t} + \sigma^R_{J_i,t}\sigma_A) + \tilde{w}_t \bar{l}_{i,t} - \tilde{c}_{i,t} + \left(-\tilde{q}_t k_{i,t} + \phi \tilde{q}_t k_{i,t} + \tfrac{\tilde{\theta}^{id}_t}{\sigma_I} \right) \varsigma_{i,t} \right] dt \\ &+ \left(\tilde{q}_t k_{i,t} \sigma_A - (\sigma^Z_{J_i,t} + \sigma^R_{J_i,t}\sigma_A) \right) dZ_t + \left(\tilde{q}_t k_{i,t} \sigma_I - \sigma^R_{J_i,t} \sigma_I \right) dZ_{i,t}. \end{aligned} \tag{E.8}$$

The term multiplying $\varsigma_{i,t}$ captures the effect of diverting capital. By diverting, the entrepreneur pays the cost $(1-\phi)\, \tilde{q}_t k_{i,t} \varsigma_{i,t}$, but benefits from lowering transfers to the principal by $\sigma^R_{J_i,t} \varsigma_{i,t}$.

Imposing $\varsigma_{i,t} = 0$ and defining

$$\tilde{\theta}^{ag}_{i,t} \equiv \sigma^Z_{J_i,t} + \sigma^R_{J_i,t}\sigma_A, \qquad \tilde{\theta}^{id}_{i,t} \equiv \sigma^R_{J_i,t}\sigma_I,$$

we obtain the law of motion of financial wealth reported in Section 3.

Transfers $\tilde{F}_{i,t}$ affect $\tilde{n}_{i,t}$ only through the diffusion terms of the principal's continuation value $\tilde{J}_i$. A convenient representation of the contract can be achieved by setting $\tilde{J}_{i,t} = 0$ and writing the contract in terms of $(\tilde{\theta}^{ag}_{i,t}, \tilde{\theta}^{id}_{i,t})$ instead of $\tilde{F}_{i,t}$. This reformulation also avoids the non-uniqueness of $\tilde{F}_{i,t}$, since an entrepreneur could, for example, borrow from the principal and invest in the risk-free asset without changing her utility.

The entrepreneur's problem. The entrepreneur's problem can be written as

$$\rho \tilde{V}_t = \max_{\tilde{c}_i, \tilde{\theta}_t^{ag}, \tilde{\theta}_t^{id}, k_i, l_i, \iota_i} \frac{c_{i,t}^{1-\gamma}}{1-\gamma} + \mathbb{E}_t[dV_t], \tag{E.9}$$

subject to the law of motion of financial wealth, the borrowing constraint $\tilde{n}_{i,t} \geq -\tilde{h}_{i,t}$, and the incentive-compatibility (IC) condition

$$0 \in \arg\max_{\varsigma_{i,t} \geq 0} \left\{ \frac{c_{i,t}^{1-\gamma}}{1-\gamma} + \mathbb{E}_t[dV_t] \right\}. \tag{E.10}$$

Applying Itô's lemma to the value function, the IC constraint becomes

$$-\tilde{q}_t k_{i,t} + \phi \tilde{q}_t k_{i,t} + \frac{\tilde{\theta}_t^{id}}{\sigma_I} \leq 0 \quad \Rightarrow \quad \tilde{\theta}_{i,t}^{id} \leq (1-\phi)\, \tilde{q}_t k_{i,t} \sigma_I, \tag{E.11}$$

where we used the fact that $V_{n,t} > 0$.

Hence, the optimal contracting problem with transfers to a principal is equivalent to problem (13), where entrepreneurs have access to aggregate and idiosyncratic insurance subject to the skin-in-the-game constraint (11).

E.3 Alternative implementation

We have shown that one implementation of the optimal contract features entrepreneurs trading in both aggregate and idiosyncratic insurance markets. This, however, is not the only way to achieve the optimal allocation. In this subsection, we consider an alternative implementation in which a financial intermediary offers a *flexible lending contract* to entrepreneurs. Unlike a standard lending contract with fixed repayments, the flexible contract allows entrepreneurs to postpone repayments after negative shocks, which are repaid following positive shocks.

For ease of exposition, we assume that intermediaries are owned by wage earners, there is no aggregate risk, and we abstract from the overlapping generations structure. These assumptions are not essential but simplify the presentation.

Financial intermediary. Let J_t denote the value of the intermediary at time t. It is the sum of entrepreneurs' obligations:

$$J_t = \int_{i \in \mathcal{E}} J_{i,t}\, di, \tag{E.12}$$

where $J_{i,t}$ is the present discounted value of repayments $F_{i,t}$:

$$J_{i,t} = \mathbb{E}_t \left[\int_t^\infty \frac{\pi_s}{\pi_t} F_{i,s} \, ds \right]. \tag{E.13}$$

Consistent with the discussion above, the dynamics of $J_{i,t}$ are

$$dJ_{i,t} = \left[r_t J_{i,t} - F_{i,t} \right] dt + \sigma^R_{J_i,t} \left(dR_{i,t} - \mathbb{E}_t[dR_{i,t}] \right).^{39} \tag{E.14}$$

By a law of large numbers (see, e.g., Sun 2006), the intermediary's portfolio J_t is not exposed to idiosyncratic risk, since such risks are diversified. As a result, loans to entrepreneurs yield the riskless rate r_t.

If $\sigma^R_{J_i,t} = 0$, we obtain a standard (riskless) lending contract, where $J_{i,t}$ evolves deterministically. If instead $\sigma^R_{J_i,t} > 0$, the contract is flexible: entrepreneurs reduce repayments after adverse shocks and increase them after favorable shocks.

Entrepreneur's financial wealth. The entrepreneur's financial wealth is defined as

$$n_{i,t} \equiv b_{i,t} + q_t k_{i,t} - J_{i,t}, \tag{E.15}$$

reflecting holdings of riskless bonds, capital, and borrowing from the intermediary.

Imposing no stealing, its law of motion is

$$dn_{i,t} = \left[r_t b_{i,t} + q_t k_{i,t} \mu^R_t - c_{i,t} + w_t \bar{l}_{i,t} - F_{i,t} - (r_t J_{i,t} - F_{i,t}) \right] dt + (q_t k_{i,t} - \sigma^R_{J_i}) \sigma_I dZ_{i,t}. \tag{E.16}$$

Note that $F_{i,t}$ cancels out of the budget constraint, and $J_{i,t}$ only enters through the diffusion term, after using the definition of financial wealth to eliminate $b_{i,t}$. Since entrepreneurs can borrow and lend freely at r_t, the value of the intermediary's contract lies in its hedging properties.

As shown above, entrepreneurs have no incentive to divert capital if $\sigma^R_{J_i,t}$ is sufficiently limited. This can be implemented through a *credit limit*. Suppose $\sigma^R_{J_i,t} = J_{i,t}$ and impose

$$J_{i,t} \leq (1 - \overline{\phi}) q_t k_{i,t}. \tag{E.17}$$

As long as $\overline{\phi} > \phi$, diversion is not optimal in equilibrium.

[39] An equivalent expression holds with aggregate risk when working with detrended variables, $J_{i,t} \equiv \tilde{J}_{i,t} / A_t$.

The entrepreneur's problem is then

$$V_t(n_i) = \max_{c_i, J_i, k_i, l_i, \iota_i} \mathbb{E}_t \left[\int_0^\infty e^{-\rho z} \frac{c_{i,t+z}^{1-\gamma}}{1-\gamma} \, dz \right], \tag{E.18}$$

subject to (12), $c_{i,t}, k_{i,t} \geq 0$, the law of motion

$$dn_{i,t} = \left[(n_{i,t} - q_t k_{i,t}) r_t + q_t k_{i,t} \mu_{i,t}^R + w_t \bar{l}_{i,t} - c_{i,t} \right] dt + \left(q_t k_{i,t} - J_{i,t} \right) \sigma_I dZ_{i,t},$$

and the credit limit (E.17), given initial wealth $n_{i,t} = n_i > -h_{i,t}$.

Similar to the skin-in-the-game constraint in the baseline model, the credit limit binds in equilibrium. A reduction in $\overline{\phi}$ can be interpreted as an expansion of credit under this lending contract, embedding insurance-like features. This flexible lending contract therefore provides an alternative institutional implementation of the optimal allocation in Section 3, where idiosyncratic insurance is delivered through state-contingent credit rather than explicit insurance markets.