WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 13-45
ENTREPRENEURIAL TAIL RISK:
IMPLICATIONS FOR EMPLOYMENT DYNAMICS

Thorsten Drautzburg
Federal Reserve Bank of Philadelphia

First Version: November 15, 2012
This Version: November 12, 2013

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Entrepreneurial Tail Risk: Implications for Employment Dynamics

First version: 11/15/2012

This version: November 12, 2013

Thorsten Drautzburg*

Abstract

New businesses are important for job creation and have contributed more than proportionally to the expansion in the 1990s and the decline of employment after the 2007 recession. This paper provides a framework for analyzing determinants of business creation in a world where new business owners are exposed to idiosyncratic risk due to initial imperfect diversification. This paper uses this framework to analyze how entrepreneurial risk has changed over time and how this has affected employment in the US. Conditions are provided under which entrepreneurial risk can be identified using micro data on the size distribution of new businesses and their exit rates. The baseline model considers both upside and downside risk. Applied to US time series data, structural estimates suggest that higher upside risk explains much of the high job creation in the late 1990s. Time variation in risk explains around 40% of the variation in employment of new businesses. Reduced form results show that this relationship is strongest in IT-related industries. When restricting the model to a single risk factor, the explanatory power for employment drops by 25% to 50% compared to the baseline estimates.

* Federal Reserve Bank of Philadelphia. E-mail: tdrautzburg[at]gmail.com. Web: https://sites.google.com/site/tdrautzburg/. I greatly appreciate the comments and guidance by my thesis committee, Lars Hansen, Tarek Hassan, Ralph Koijen, and particularly my committee chair, Harald Uhlig. I also benefited from comments by George Alessandria, Enghin Atalay, Philip Barrett, Jaroslav Borovicka, Lasse Brune, John Cochrane, Pablo Guerron-Quintana, Erik Hurst, Bryan Kelly, Erik Loualiche, Ezra Oberfield, Monika Piazzesi, Laura Pilossoph, Jon Willis, Mu-Jeung Yang, and seminar participants at the Chicago, Kansas City, New York, Philadelphia, Richmond, and the St. Louis Feds, the FRB, UIUC, UCSD, the University of Chicago, Vanderbilt University, and the SED 2013 meetings. I am grateful for the hospitality of the Kansas City Fed during an early stage of this paper. All remaining errors are mine. The views expressed herein are those of the author. They do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or its Board of Governors. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/.

1 Introduction

What role do young businesses play for employment dynamics? In the US, net job creation by new businesses exceeds that of the economy as a whole, though many of these jobs are subsequently destroyed. Surviving young businesses grow, on average, more quickly than existing businesses (Haltiwanger et al., 2010). New businesses contributed more than existing businesses to the private sector employment growth in the late 1990s, but they also contributed more than proportionally to the decline in employment from 2007 to 2010. Since entrepreneurs can diversify only imperfectly (Moskowitz and Vissing-Jørgensen, 2002; Hall and Woodward, 2010), a natural question to ask is whether changes in entrepreneurial risk are important for understanding these aggregate changes of employment at young businesses.

This paper provides a tractable dynamics macro model for analyzing entrepreneurial risk and its effect on business creation and employment. In my model, risk-averse individuals, who differ in their known skills and face idiosyncratic productivity risk, make entry and hiring decisions. Due to limited diversification, the composition of idiosyncratic entrepreneurial risk is crucial for them. Despite the heterogeneity of firms and heavy-tailed productivity distributions, an aggregation result of the entrepreneurial sector keeps the model tractable, allowing me to solve for dynamics quasi-analytically in a special case. The model connects, moreover, in a natural way with the data: Conditions are provided under which time variation in entrepreneurial risk is identified semi-structurally from publicly available data on the cross-section of new businesses. These identified shocks are then used as an input to estimate a fully stochastic version of the model with capital adjustment costs and wage rigidity.

Figure 1 illustrates the changing contribution of young businesses to employment in the US. Businesses up to one year of age make up, on average, 10% of the total private sector employment. From 1994 to 2000, employment at young businesses up to one year old increased by about 2 million jobs, or 1.2% of the working age population, which is about one quarter of the total aggregate increase. From 2007 to 2010, employment at young businesses declined by 1.8% of the working age population (left panel), about 35% of the total decrease. These more than proportional changes are reflected in an overall employment share that is first rising and then falling over these periods (right panel).[1]

I model the changing employment contributions of young businesses as driven by entry, exit, and hiring decisions of young entrepreneurs. These entrepreneurs are a self-selected subset of newborn agents in a Blanchard (1985)-type perpetual youth model.

[1]Here, businesses are defined as establishments, and job creation is employment at young establishments. Figure C.12 in the appendix shows analogous results with alternative measures for businesses (namely as firms) and genuine net job creation.

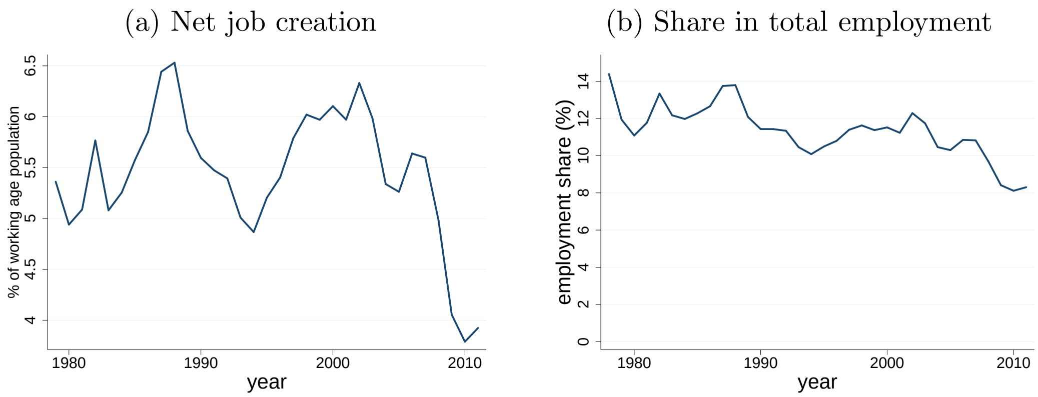


Figure 1: Young businesses: net job creation and employment share, 1979–2011

Note: Young businesses rapidly grow to a large share of the US economy, but were disproportionately affected by the recent recession (right panel). New businesses create on average new jobs for 5% of the working age population, or 12 million people in terms of the 2010 population. Net job creation here is defined as of paid employees by businesses aged ≤ 1 year. The share is relative to total employment at all businesses, which mostly excludes the public sector. All data are from the Business Dynamics Statistics at the establishment-level.

Elastic labor supply is introduced by allowing for an additional extensive margin between home-production and market work. Initial entrepreneurial business income risk and subsequent diversification are modeled in two stages. In the beginning of the first period, newborn agents cannot diversify, for example, due to limited commitment. After entry and exit decisions are made, all agents have access to complete markets so that idiosyncratic risk is fully diversified in equilibrium. This incomplete-complete markets setup is one key ingredient to keep the model tractable. The other key ingredient is the abstraction from further selection and dynamics among existing businesses as in Hopenhayn (1992). Instead, exit of established businesses is exogenous and their expected growth rate independent of size and age.[2] Businesses hire factor inputs on competitive spot markets. Size is determined as in Lucas (1978) through a limited span of control. These decreasing returns together with competitive spot markets allow me to model the entrepreneurial sector as a representative firm with an aggregate span of control. This also allows me to easily value existing businesses and is the second key ingredient to keep the model tractable. Incorporating the characterization of established businesses in the problem of newborn agents, I show that entry depends only on idiosyncratic risk and on the state of the economy through relative present discounted profits as a sufficient statistic. Under certain distributional assumptions, the exit rate is shown to be independent

[2] It is straightforward to allow for common cohort-life-cycle changes in average growth rates. In a partial-equilibrium Hopenhayn model consistent with life-cycle dynamics, Clementi and Palazzo (2013) find that life-cycle dynamics propagate aggregate shocks. However, Clementi et al. (2013) indicate that these results may not survive in general equilibrium.

of macroeconomic conditions other than entrepreneurial risk.[3]

To assess the empirical hypothesis that time-varying entrepreneurial risk can explain the employment dynamics of young businesses, I distinguish between upside risk and downside risk. These correspond to the fatness of the left and right tail of the productivity distribution in my model.[4] Together, both risk components capture the salient feature of a skewed size distribution and high exit rates among new entrepreneurs. To recover them from the data, I make use of an estimator derived in Welsh (1986) for the Hall (1982) class of distribution that asymptote to a power law. Since I show that the right tail of the size distribution of young businesses is approximately Pareto, this estimator allows me to construct a time series of upside risk from repeated cross-sections of young businesses. To deduce ex ante risk from the estimated tail coefficient, I assume that the difference in size among the largest businesses is primarily driven by risky productivity draws. Entrepreneurial skills also matter, but they are second order when conditioning on the largest businesses. My time series for downside risk is based on exit rates and the parametric assumptions of my model, but it can be computed knowing only upside risk and exit rates. Next, I construct a tractable general equilibrium model, in which the risk exposure is of first order to new entrepreneurs and which incorporates the distributional shapes I find in the micro data. The model parametrically identifies downside risk from exit rates. Identifying these risk shocks from the cross-sectional information of different cohorts of young businesses, I provide a structural Bayesian estimation of my model. This structural estimation allows me to decompose observed employment at young businesses into different components. I find that about 40% of the employment creation variation by young firms is due to variations in upside and downside risk. Combining both risk components into a common risk factor lowers the explanatory power by 30% or more.

To validate my estimates, I consider additional data. I find that the model qualitatively fits data on entry and average size that are not used in the estimation. It also fits the cross-sectional size distribution well, even though only the upper 1% is used in the baseline estimation. Using a reduced form unrestricted VAR, I also find suggestive evidence to support my model at the sector level. I find that the predictions for employment, entry, and average size are most robust in SIC industries 70–89 (services) and, for establishments, in industries 40–49 (transportation and utilities). These sectors include telecommunications and software development, respectively, suggesting that increased upside risk in my model may reflect new technologies. In line with my structural es-

[3] The independence of the exit rate from macroeconomic conditions is driven by offsetting selection effects and exit incentives when productivity follows a power law. To the extent that this would hold approximately under more general distributions, this explains the lack of cyclicality of exit rates observed in Lee and Mukoyama (2008).

[4] Productivity in my model is a convolution of Double-Pareto distributions used in, for example, Reed and Jorgensen (2004) and Arkolakis (2011).

timates, I also find that entry is mostly driven by downside risk, while changes in the average firm size are driven by both upside and downside risk.

The remainder of the paper is organized into three parts. First, I provide a review of related literature. Second, I outline the model and analyze its properties. Third, I estimate my model for US data. The appendix contains proofs and robustness checks.

2 Related Literature

The importance of young or small business for employment has been the subject of a growing literature. Moscarini and Postel-Vinay (2009) analyze a dynamic version of the Burdett-Mortensen search model and point out implications for differential growth of small and large firms over the business cycle. Their mechanism is absent from my model. Instead, I include a reduced form wage rigidity as in Blanchard and Galí (2007) and Uhlig (2007) for quantitative purposes, which affects all employers alike. Cooley et al. (2004) analyze an optimal contracting problem of an entrepreneur with limited commitment in general equilibrium and find that because young entrepreneurs have less collateral to pledge, young and small firms grow faster than big firms that have reached their unconstrained optimal scale. The stylized, initially limited but eventual full commitment I use to derive the initial lack of risk-sharing in Appendix C.2 can be viewed as continuing entrepreneurs growing out of financial frictions. In a numerical extension (Appendix C.5), I also show that the introduction of risky debt mostly affects the macroeconomic impact of downside risk shocks in my model. Unlike Cooley et al. (2004), the choice to become an entrepreneur in my model is endogenous. Endogenous and time-varying entry and exit also differentiate my paper from models such as Veracierto (2008) and Hopenhayn and Rogerson (1993) and should be viewed as a complement to their focus on firm dynamics, which is absent from my paper. Clementi and Palazzo (2013) also feature an endogenous number of businesses, but only in partial equilibrium.

As reported by Haltiwanger et al. (2010), although small businesses in general have often been perceived as being key for private sector employment creation, the important characteristic is young age and not small size. The modeling approach in my paper builds on this idea. It is also in line with Hurst and Pugsley (2011), who caution against viewing all small businesses as potential growth engines since nonpecuniary motives drive many entrepreneurs. I address this issue by distinguishing not only entrepreneurs and workers, but additionally the self-employed. Unlike entrepreneurs, the self-employed have a constant scale technology in my model. My focus on young businesses is also motivated by Jovanovic and Rousseau (2002) who show that many of the most successful US businesses in history had their fundamental innovations early in their lifecycles.

The premise of this paper that entrepreneurs are particularly exposed to idiosyncratic risk is well documented. Bradford and Sokolyk (2012) find that new business owners typically invest around 50% of their net worth in their businesses. Business income risks are substantial and much larger than for labor income, as documented by DeBacker et al. (2012) for US tax data. Hall and Woodward (2010) also document that risk is resolved early in firms' lifecycles, particularly for successful businesses. Value-weighted, about three quarters of the venture capital-funded entrepreneurs in their data either sell off or exit within five years. This motivates the empirical specification of risk as being front-loaded in this model.

My paper provides a link between the literature on entrepreneurial risk and the literature on time-varying risk. The seminal paper in the literature is the structural model by Bloom (2009), but additional empirical studies documenting time variation in idiosyncratic risk include Campbell et al. (2001), and Bloom et al. (2011). Reduced form empirical analyses have shown that the processes driving investment risk contain multiple components (e.g., Adrian and Rosenberg, 2008). My paper distinguishes two types of idiosyncratic risk for for entrepreneurs, namely upside and downside risk.

Measuring entrepreneurial risk is challenging. Moskowitz and Vissing-Jørgensen (2002) look at firm exit rates and private equity returns conditional on firm survival for entrepreneurial households, using cross-sectional data from the triennial Survey of Consumer Finances from 1989 to 1998. They document high exit rates and a wide dispersion of returns conditional on survival, but they do not analyze time variation in risk. An important recent contribution by DeBacker et al. (2012) uses tax data to quantify business income risk. This is complementary to the approach taken here. I use repeated cross-sections of establishment-level data to back out time-varying risk from the size distribution and the exit rate of establishments. This approach is similar to Kelly (2012), who uses the cross-sectional variation to estimate common tail risk of financial returns.

Other mechanisms besides risk aversion and market incompleteness have similar predictions. Instead of risk aversion paired with undiversifiable risk, one could consider investment adjustment costs to match responses to changes in risk (e.g., Bloom, 2009; Bloom et al., 2011; Lee, 2011). Both uninsurable risk and risk aversion imply a technological exposure to idiosyncratic risk. Indeed, entrepreneurial risk aversion in this paper is akin to adjustment costs in reducing the option value of investment opportunities. Besides risk aversion and costly adjustment, also a concern for robustness as in Hansen and Sargent (2007) is a justification for qualitatively similar investment behavior in a model with risk-neutral investors.

This paper is a neoclassical growth model with intangible capital (Hall, 2001). Here, intangible capital is the total productivity of entrepreneurs, similar to Bilbiie et al. (2012).

Similar to McGrattan and Prescott (2010), I find that intangible capital in my model accounts for much of the boom in the 1990s. Intangible capital in my model arises from the selection between entrepreneurs, as in Luttmer (2007).

Throughout this paper, risk is taken as a primitive. However, several recent papers point out that risk may be an endogenous outcome, for example when agents experiment. Pastor and Veronesi (2009) is empirically related insofar as their model rationalizes the IT boom in the 1990s. In their model, agents choose to experiment with a new technology, causing a transient increase in uncertainty. While their endogenous learning mechanism is outside of my model, my empirical results are consistent with their idea that underlying the 1990s boom was uncertainty regarding new technology.

3 Model

The economy is inhabited by a continuum of agents, a fraction of which die every period and are replaced by new agents. At the heart of the model is a Roy-type occupational choice problem that new agents face. They choose among entrepreneurship, self-employment, and salaried work. New entrepreneurs initially face uninsurable income risk. Workers and the self-employed face no idiosyncratic income risk.[5] Growth of established businesses is *iid* in the cross-section, and their owners face complete markets.

3.1 Model setup

At any point in time t, there is a unit measure of households i in the set $f\mathcal{I}_t$. I use the subscript iat for household i of age a at time t. Households have preferences over the stream of their lifetime consumption $\{C_{iat}\}$. Each household lives for a random number of years, dying with an *iid* probability θ. At the same rate, new households are born, keeping the measure of households constant over time.[6]Newborn agents observe the aggregate state as well as an idiosyncratic signal about their future productivity. Furthermore, they are randomly chosen to be a member of two fractions of the population. A fraction ϵ of agents gets to choose between becoming an entrepreneur or a worker. The remaining fraction $1-\epsilon$ chooses between becoming self-employed or a worker. After they choose, their idiosyncratic component of future productivity is fully realized. While agents are initially fully exposed to this risk due to limited commitment before decisions are made, they face complete markets after choosing entry and exit.

[5] I could relax this assumption as long as nonentrepreneurs face less risk than entrepreneurs. In the data, business income is much riskier than labor income (cf. DeBacker et al., 2012).

[6] Formally, as pointed out by Judd (1985) and discussed further by Uhlig (1996), if households die at an *iid* rate, the integral over all households of an indicator function indicating survival is not defined. Following the literature I assume that a law of large numbers holds, despite the technical issues.

3.1.1 Occupations

Entrepreneurs hire capital and labor to produce output, subject to decreasing returns to scale. The self-employed operate a fixed-scale technology. Workers differ in their efficiency units of labor and are otherwise homogeneous. Agents' skills are summarized by their productivity Z_{iat} and labor endowment H_{iat}.

An entrepreneur with productivity parameter Z_{iat} hires capital and labor K_{iat}, N_{iat} to produce output Y_{iat}, according to the following production function:

$$Y_{iat} = AZ_{iat}^{1-\phi}\left(K_{iat}^{\alpha}N_{iat}^{1-\alpha}\right)^{\phi}, \tag{3.1}$$

where $1-\phi$ is the profit share and α is the cost share of capital. Productivity Z_{iat} is scaled by $1-\phi$ so that profits and the policy function are linear in Z_{iat}. $A > 0$ is a constant used for choosing units in the empirical section.

The fixed-scale technology of the self-employed is simple: Each self-employed only uses her own labor to produce an output equal to her productivity:

$$Y_{iat} = Z_{iat}. \tag{3.2}$$

Workers are homogeneous, except for differences in their endowment with efficiency units of labor, H_{iat}.

3.1.2 Information and type distributions

At the beginning of a period, all newborn households observe an idiosyncratic, type-dependent random variable O_{i0t} as well as the aggregate state of the economy ζ_t. At that time, they choose their occupation. The households in the fraction $1-\epsilon$, which choose between self-employment and being a worker, will have productivity given by:

$$Z_{i0t} = O_{i0t}\omega_t \tag{3.3}$$

when self-employed, where ω_t is an aggregate and exogenous parameter, varying with t and with growth rate $g_{s,t} = \frac{\omega_t}{\omega_{t-1}}$. They have human capital $H_{iat} = 1$, when deciding to become workers. Their occupational choice is fully reversible.

The fraction ϵ of households, which chooses between becoming an entrepreneur and being a worker, will have a productivity parameter given by:

$$Z_{i0t} = O_{i0t}U_{i0t} \tag{3.4}$$

when choosing to become an entrepreneur, where U_{i0t} is an additional idiosyncratic random variable, directly revealed after their occupational choice. Entrepreneurial productivity growth is defined as $g_{e,iat} = \frac{Z_{iat}}{Z_{i,a-1,t-1}}$. They will have human capital $H_{iat} = 1$, when deciding to become a worker at the beginning of the period. If they exit their business during the first period of their lives, they have only human capital $H_{iat} = 1 - \eta$.

For the components of initial entrepreneurial productivity, I assume independent distributions with heavy tails. A parsimonious way to introduce a potentially asymmetric distribution with heavy tails is to model each tail of the distribution as a Pareto distribution. Imposing continuity for the density function yields the Double-Pareto distribution. Equivalently, the logarithm of risky productivity U_{i0t} follows a Laplace distribution:[7]

$$\ln U_{i0t} \overset{iid}{\sim} \mathcal{L}(\ell_t, r_t) + \mu_t, \tag{3.5}$$

where $\mathcal{L}(\ell, r)$ denotes the centered, asymmetric Laplace distribution with left and right tail parameters ℓ and r. μ_t is a location parameter. I use the same distributional form for the observed type of entrants:

$$\ln O_{i0t} \overset{iid}{\sim} \mathcal{L}(\lambda, \nu).$$

The *iid* assumption means that O_{i0t} and U_{i0t} are independently distributed within any given period t, but the time variation in the distribution implies that they are not identically distributed across periods.

Figure 2 illustrates the distribution of potential productivity. The dashed and dotted lines represent the densities of the initially unknown component $\ln U$ and the observed component $\ln O$. The solid line is the convolution of these two densities and represents the density of potential log productivity $\ln(U \times O)$, prior to self-selection of entrepreneurs. The densities are drawn asymmetrically in anticipation of the empirically relevant parameter range.

The skill of the self-employed is also independently distributed in the cross-section, drawn from a time-invariant distribution F_O:[8]

$$O_{i0t} \overset{iid}{\sim} F_O. \tag{3.6}$$

For ease of notation, I abstract from separate time indices for the beginning and the end of the period. I use $\mathcal{F}_{i0t}^{-}$ to denote the beginning of the period information set, when

[7] The Laplace density is given by: $f_{\mathcal{L}(\ell,r)}(x) = \frac{1}{\ell^{-1}+r^{-1}} \min\left\{e^{\ell x}, e^{-rx}\right\}$. For a derivation and description of the Double-Pareto distribution see Reed and Jorgensen (2004); Arkolakis (2011). They consider a log-normal Laplace distribution. To simplify the characterization of the exit decision, I set the normal component to zero. Most qualitative results in this section extend to the log-normal Laplace case.

[8] While nothing changes conceptually when the distribution of O_{i0t} is the same as for entrepreneurs, a calibrated version implies very thin tails of F_O in this case. I therefore use a log-normal distribution.

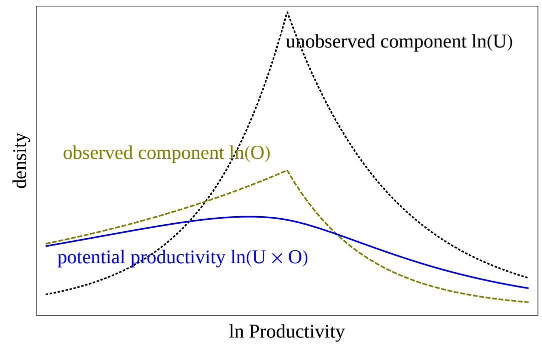


Figure 2: Density of potential productivity and its observed and unobserved component
Potential productivity combines observed skill O and productivity risk U. Productivity risk has a fatter right tail than the skill distribution. The dotted line is the density of the unobserved component $\ln U$, the brown dashed line the density of the observed component $\ln O$ and the blue solid line the density of potential $\ln Z$ prior to selection into entry.

U_{i0t} is unknown, but the aggregate state ζ_t is already known. I denote the history of ζ_t as ζ^t. Under the end of the period information set $\mathcal{F}_t$ also the individual states are perfectly observed by all agents.

3.1.3 Preferences

The preference specification follows Garleanu and Panageas (2010) in combining the perpetual youth model of Blanchard (1985) with the recursive utilsity framework in Epstein and Zin (1989) and Weil (1990). Agents care only about consumption $\{C_{iat}\}$.[9]

Given a path of consumption $\{C_s\}_{s=t}^{\infty}$, household i of age a has value $\tilde{V}$:

$$\tilde{V}_{iat}(\{C_s\}_{s=t}^{\infty}) = \Big((1-\beta)(C_t)^{1-\psi} + \beta\mathcal{R}[\tilde{V}_{i,a+1,t+1}(\{C_s\}_{s=t+1}^{\infty})|\mathcal{F}_{iat}]^{1-\psi}\Big)^{\frac{1}{1-\psi}}, \qquad (3.7)$$

where $\mathcal{R}$ denotes the following risk-adjusted expectation operator:

$$\mathcal{R}[\zeta|\mathcal{F}] \equiv \mathbb{E}[\zeta^{1-\rho}|\mathcal{F}]^{\frac{1}{1-\rho}}.$$

Agents do not consume within the period. The beginning of the period value $\tilde{V}_{i0t}^{-}$ is therefore just the risk-adjusted continuation value:

$$\tilde{V}_{i0t}^{-}(\{C_s\}_{s=t}^{\infty}) = \Big(0 + \beta\mathcal{R}[\tilde{V}_{i,a+1,t+1}(\{C_s\}_{s=t+1}^{\infty})|\mathcal{F}_{i0t}^{-}]^{1-\psi}\Big)^{\frac{1}{1-\psi}}$$

[9]The preferences can also be derived using a concern for robustness with two different uncertainty adjustments for survival and investment risk in the framework of multiplier preferences in Hansen and Sargent (2007). They allow for different concerns about misspecification within and across Markov regimes. Here, the analogues are different concerns for misspecification of the survival probability versus the wealth evolution. See Appendix C.1.

$$= \beta^{\frac{1}{1-\psi}} \mathcal{R}[\tilde{V}_{i,a+1,t+1}(\{C_s\}_{s=t+1}^{\infty})|\mathcal{F}_{i0t}^{-}]. \tag{3.8}$$

This specification implies that households have an intertemporal elasticity of substitution of ψ^{-1} along a deterministic consumption path. ψ is also the aversion over mortality risk, which affects a household with *iid* probability θ. In contrast, households adjust investment risk according to the risk aversion parameter ρ. The risk-adjusted expectation operator $\mathcal{R}$ only incorporates this investment risk. $\beta = \tilde{\beta}(1-\theta)$ reflects both the rate of time preference $\tilde{\beta}^{-1} - 1 > 0$ and the death probability $\theta \in (0, 1)$.

3.1.4 Risk sharing

Actuarial fair insurance is available to agents in this economy, except at birth. From the end of the first period onward, agents have access to complete markets, following Blanchard (1985). Actuarial fair insurance combined with risk aversion implies that agents completely insure against idiosyncratic risks from the end of the first period onward. Risk prices for idiosyncratic risk simply reflect physical probabilities. At birth, however, agents cannot insure against idiosyncratic risk. Appendix C.2 explains this with an initial lack of commitment and limited information in an environment with perfect competition among risk-neutral lenders.

The appendix also shows that income from financial assets is discounted at a different rate than income from human capital. Since human capital is wiped out at death, but financial assets can be passed on to surviving agents, the discount factor for financial asset income is $(1-\theta)^{-1}$ times the discount factor for human capital income.

3.1.5 Capital goods sector

There is a competitive sector supplying aggregate capital. Aggregate supply of capital follows a standard law of motion with a stochastic depreciation rate. Investment is subject to adjustment costs in the investment rate:

$$K_t = (1-\delta_t)K_{t-1} + I_t\left(1 - \Psi(I_t/K_{t-1})\right), \tag{3.9}$$

where $\Psi(x) = \frac{\bar{\Psi}}{2}\left(x\left(\frac{\bar{I}}{\bar{K}}\right)^{-1} - 1\right)^2, \bar{\Psi} > 0$.

Dividends in the capital goods sector are given by $d_t^s = K_t d_t^k - I_t$. The variable $\Pi_{k,t}$ denotes their present discounted value:

$$\Pi_{k,t} = \sum_{t=0}^{\infty} \int_{\zeta^t} \frac{Q_t(\zeta^t)}{Q_0(1-\theta)^t} \Big(K_t(\zeta^t)d_t^k(\zeta^t) - I_t(\zeta^t)\Big) d\zeta^t. \tag{3.10}$$

3.1.6 Labor market

The labor market is competitive. Entrepreneurs can hire one efficiency unit of labor at the given wage rate W_t.

To fit employment movements by existing firms, I consider a friction in the labor market that causes wages to be rigid in the short run. Real rigidities are often proposed to explain aggregate labor market outcomes (e.g., Hall (2005) and Shimer (2005)). Here, I follow Blanchard and Galí (2007) and Uhlig (2007). I assume that real wages paid to workers adjust only partially to changes in labor demand. In particular, while firms pay a wage rate W_t, workers receive W_t^f. The difference is a time-varying labor-wedge:

$$W_t = (W_t^f)^{1-\kappa} W_{t-1}^{\kappa}, \tag{3.11}$$

where $\kappa \in [0, 1)$ is the persistence of real wages.

3.1.7 Budget constraint

Before agents start consuming, they have access to complete markets and can trade Arrow-Debreu securities $B(\zeta_i^t)$, which pay one unit of consumption in history ζ_i^t at price $Q(\zeta_i^t)$. Markets reopen every period t, and $B_{i,a-1,t-1}(\zeta_j^s) = b$ denotes purchases of b consumption claims in history ζ_j^s by agent i at time $t-1$ at age $a-1$. Additionally, they hold shares in the capital goods producing firms, A_{iat}, priced at P_t^s and paying a dividend of d_t^s.

To describe the budget set, I make use of two properties that hold in any competitive equilibrium. First, since insurance is fair, agents insure completely against idiosyncratic risk, other than the effect of mortality on human capital. Second, since factor markets are competitive and entrepreneurs hire capital and labor after observing their productivity, they maximize profits state by state. I show in Appendix A.1 that the resulting profits are linear in entrepreneurial productivity in the cross-section:

$$Y_{iat} - N_{iat} W_t - K_{iat} d_t^k = \bar{\Pi}_t Z_{iat}, \tag{3.12}$$

where $\bar{\Pi}_t$ is a function of the aggregate state. Since they insure against idiosyncratic growth risk after their initial productivity draw is realized, their budget depends on $Z_{iat} \equiv Z_{i,0,t-a} G_{e,t-a,t}$. $G_{e,t-a,t}$ is the average productivity growth in the entrepreneurial sector.[10]

Let π_{iat} denote the period profits agent iat's human capital generates in period t,

[10]Formally, $G_{e,t-a,t} = \prod_{\tau=t-a+1}^{t} g_{e,\tau}$, where $g_{e,t} \equiv \mathbb{E}[g_{iat} | \mathcal{F}_t^-] = \mathbb{E}\left[\frac{Z_{iat}}{Z_{i,a-1,t-1}} | \mathcal{F}_t^-\right]$.

with wages understood as the profits derived from a worker's labor endowment. For the different occupational choices ι_{iat}, profits are given as follows:

$$\pi_{iat} = \begin{Bmatrix} O_{i0t}\bar{w}_t & \iota_{iat} = se \\ W_t^f & \iota_{iat} = n \\ (1-\eta)W_t^f & \iota_{iat} = x \\ \bar{\Pi}_t G_{e,t-a,t} Z_{i0t} & \iota_{iat} = e \end{Bmatrix},$$

where ι_{it} denotes the job choice of agent i, with s representing the self-employed, n salaried work, χ exited entrepreneurs, and e active entrepreneurs.

Using this notation, the budget constraint can be expressed as follows:

$$\begin{aligned}\sum_{t=0}^{\infty} \int_{\zeta^t} \frac{Q_t(\zeta^t)}{Q_0(1-\theta)^t} \Big(& B_{i,a-1,t-1}(\cdot) + A_{i,a-1,t-1}(\cdot)(P_t^s(\cdot) + d_t^s(\cdot)) \\ & - (P_t^s(\cdot)A_{iat}(\cdot) + B_{iat}(\cdot))\Big) d\zeta^t \\ \geq \sum_{t=0}^{\infty} \int_{\zeta^t} \frac{Q_t(\zeta^t)}{Q_0} & \Big(C_{iat}(\zeta^t) - \pi_{iat}(\zeta^t)\Big) d\zeta^t \end{aligned} \tag{3.13}$$

with $B_{i,a-1,t-1}(\cdot)$ and $A_{i,a-1,t-1}(\cdot)$ given. Young agents have no endowments of Arrow-Debreu securities and stocks: $B_{i,-1,t-1}(\cdot) = A_{i,-1,t-1} = 0$.[11] Thus, for the young equation (3.13) simply bounds the present discounted value of consumption by the present discounted value of lifetime human capital income.

3.2 Equilibrium

3.2.1 Household optimization problem

Figure 3 summarizes the key occupational choice problem faced by agents in this model. Each occupational choice has a lifetime value attached to it, and households choose the career path with the highest value. Two features are crucial. First, entry is costly for entrepreneurs so that not all potential entrepreneurs are willing to enter. Second, nonentrepreneurs can switch seamlessly back and forth between salaried work and self-employment, providing an elastic supply of labor in equilibrium.[12]

Conditional on their occupational choice between occupations ι_i, households choose consumption, net purchases of Arrow-Debreu securities $B_{iat}(\cdot)$, and shares in the capital

[11] In a numerical extension I introduce positive endowment of financial wealth, along with a pecuniary cost of becoming an entrepreneur.

[12] The assumption of two distinct subsets of the population choosing between self-employment and entrepreneurship can be relaxed. A particularly simple case arises when every agent has the same productivity when self-employed. Alternatively, additional ex ante heterogeneity can explain the different fractions of the population.

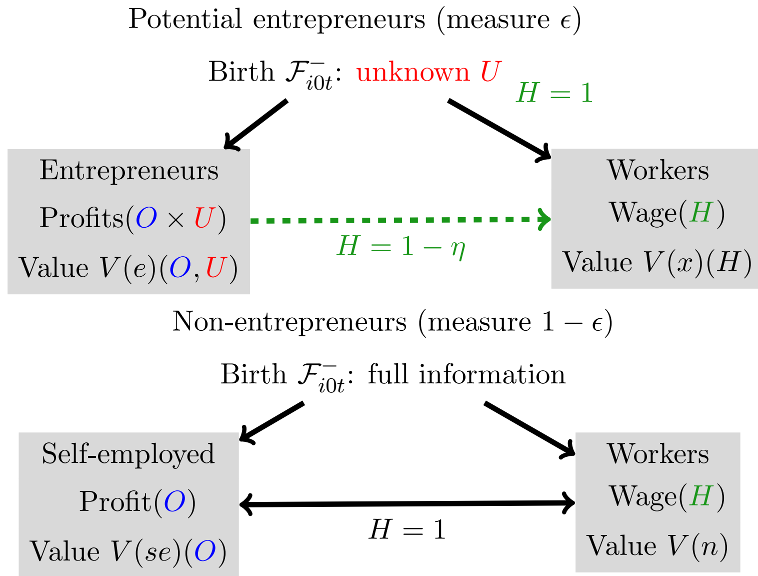


Figure 3: Occupational choices, associated period incomes and lifetime values
Agents face the choice between two occupations. Nonentrepreneurs can switch costlessly between self-employment and salaried work. Entrepreneurs who enter forgo a fraction η of their labor endowment as workers. Choices depend only on income streams: Agents have no preferences over occupations. First, potential entrepreneurs decide about entry, knowing the aggregate state and their skill. After entry, productivity U is revealed and entrepreneurs can exit, prior to production in the first period.

goods sector A_{iat} freely, given state prices $Q_t(\zeta_i^t)$, wages w_t, capital prices and rental rates P_t^k, d_t^k, and their information set $\mathcal{F}_{it}$. They solve

$$V_{iat}(\iota) \equiv \max_{\{C_{i,a+\tau,t+\tau}(\cdot), A_{i,a+\tau,t+\tau}(\cdot), B_{i,a+\tau,t+\tau}(\cdot)\}_{\tau=0}^{\infty}} \tilde{V}_{iat}(\{C_{i,a+\tau,t+\tau}(\cdot)\}_{\tau=0}^{\infty}), \tag{3.14}$$

subject to the budget constraint (3.13), given $\iota_{it} = \iota$ and zero initial financial assets. The utility function $\tilde{V}_{iat}$ here is defined in (3.7). $V_{iat}(\iota)$ is the value, in utility terms, of household ia at time t choosing $\iota_i = \iota$.

The occupational choice between entrepreneurship ($\iota = e$) and salaried work ($\iota = w$) for the measure ϵ of potential entrepreneurs is subject to information set at the beginning of period 0, $\mathcal{F}^-_{i0t}$ for household $i0$ born at time t. Because it concerns purely investment risk that is risk-adjusted according to ρ, the occupational choice is between:

$$\hat{V}_{iat} \equiv \max\{\mathcal{R}[\max\{V_{iat}(e), (1-\eta)V_{iat}(w)\}|\mathcal{F}^-_{i0t}], \mathcal{R}[V_{iat}(w)|\mathcal{F}^-_{i0t}]\}, \tag{3.15}$$

where the optimal exit decision is already incorporated for entrepreneurs. If an entrepreneur exits, $\iota = x$ and I used that the lifetime value will be reduced proportionally by the exit cost in equilibrium: $V_{iat}(x) = (1-\eta)V_{iat}(w)$.

The potentially self-employed choose between self-employment $\iota_{it} = se$ and salaried work $\iota_{it} = n$ period by period under information $\mathcal{F}^-_{it}$:

$$\hat{V}_{iat} \equiv \max\{\mathcal{R}[V_{iat}(se)|\mathcal{F}^-_{i0t}], \mathcal{R}[V_{iat}(n)|\mathcal{F}^-_{i0t}]\}. \tag{3.16}$$

Entrepreneurs who continue to operate choose state-dependent labor demand and capital demand to maximize the present discounted value of their profits. Equivalently, they maximize profits state by state:

$$\max_{N_{iat}, K_{iat}} AZ^{1-\phi}_{iat}(K^{\alpha}_{iat}N^{1-\alpha}_{iat})^{\phi} - N_{iat}W_t - K_{iat}d^k_t. \tag{3.17}$$

For notational purposes, set production and input demand by households that do not operate a firm to zero: $Y_{iat} = K_{iat} = N_{iat} = 0$ if $\iota_i \neq e$. Similarly, $H_{iat} = 0$ for entrepreneurs.

3.2.2 Capital good producers

The representative capital goods producer chooses investment and capital to maximize the expected discounted value of period profits $\Pi_{k,t}$ in (3.10), subject to the law of motion for capital (3.9), given initial capital. P^k_t denotes the current value multiplier on the constraint.

3.2.3 Competitive equilibrium

The following defines a competitive equilibrium in this economy:

Definition 1. *An allocation* $\{\{K_t, I_t, \{\iota_{iat}, C_{iat}, A_{iat}, B_{iat}, N_{iat}, K_{iat}\}_{iat}\}_{i\in\mathcal{I}_t}\}^{\infty}_{t=0}$ *and state prices, capital prices, share prices, and wages* $\{Q_t(\cdot), P^k_t, P^s_t, W_t, W^f_t\}$ *are a competitive equilibrium if, for any history and initial conditions* $\{\zeta^t_i\}_{it}$*, the allocation solves each household's maximization problem defined by* (3.14)*,* (3.15) *or* (3.16)*, the profit maximization problem* (3.17)*, and the capital goods producer's problem of maximizing* (3.10)*. In addition, the following market clearing conditions hold: (1) Goods market clearing:* $\int_{\mathcal{I}_t}(C_{iat} - Y_{iat})di + I_t = 0$*; (2) labor market clearing:* $\int_{\mathcal{I}_t}(H_{iat} - N_{iat})di = 0$*; (3) capital market clearing:* $\int_{\mathcal{I}_t} K_{iat}di = K_t$*; (4) stock market clearing:* $\int_{\cup_{s\leq t}\mathcal{I}_s} A_{iat}di = 1$*; (5) security market clearing:* $\int_{\cup_{s\leq t}\mathcal{I}_s} B_{iat}di = 0$.

A particular equilibrium without macro shocks greatly facilitates the analysis of the general model, because aggregates are constant, but individual allocations are non-degenerate. Therefore, such an equilibrium is called the stationary equilibrium:

Definition 2. *A stationary equilibrium is such that the aggregate state of the economy is constant:* $\ell_t = \bar{\ell}, r_t = \bar{r}_t, \delta_t = \bar{\delta}, \omega_t = \bar{\omega}, \mu_t = \bar{\mu}$.

3.3 Properties of the equilibrium

Proposition 1 establishes three general properties of this economy under regularity conditions. First, there is a sorting equilibrium. Second, there is a representative consumer. Third, the entrepreneurial production function aggregates to a neoclassical production function with endogenous productivity that follows a simple autoregressive process.

The regularity conditions require the tails of the productivity distribution not to be too heavy:

Assumption 1. *Assume the following conditions are satisfied for all* ζ^t*:*

1. *The right tail coefficient* r_t *satisfies* $1 < r_t < \nu$.
2. *If exit is infinitely costly,* $\eta = 1$*, the left tail coefficient* ℓ_t *satisfies* $(\rho - 1) < \ell_t$.

In addition, I assume that productivity growth rates are *iid* in the cross-section and shock-processes are (trend-)stationary.[13]

Assumption 2. *The growth rate of entrepreneurial productivity is iid in the cross-section with mean* $\bar{g}_e = \mathbb{E}_t[g_{iat}]$. $r_t, \ell_t, \delta_t, \mu_t - t\ln(\bar{g}_e)$*, and* $\omega_t - t\ln(\bar{g}_s)$ *are stationary stochastic processes with* $\bar{g}_e^{\frac{1-\phi}{1-\alpha\phi}}, \bar{g}_s < \tilde{\beta}^{-1}$.

For the sorting solution of nonentrepreneurs to have a nondegenerate solution requires an interior equilibrium in the labor market: Some, but not all nonentrepreneurs are self-employed.[14] Under this assumption, the following proposition holds:

Proposition 1. *Under Assumptions 1 and 2, the following properties hold in any interior equilibrium:*

(a) Sorting equilibrium: The occupational choice problems of potential entrepreneurs and self-employed have cutoff solutions with entry cutoffs $\bar{O}_{e,t}, \bar{O}_{se,t}$*, and an exit cutoff* $\underline{Z}_{e,t}$ *for entrepreneurs.*

[13] The model can be generalized to include life-cycle productivity dynamics, while remaining tractable when assuming that productivity growth depends only age, but not on size.

[14] Formally, this condition requires that at the equilibrium wage W_t^f: $0 < F_O(W_t^f/\omega_t^-) < 1$. I calibrate the mean productivity of entrepreneurs $\bar{\mu}$ and the wage rate per efficiency unit of labor such that the equilibrium is interior. If aggregate shocks are sufficiently small and bounded, any equilibrium is therefore interior. Conversely, an interior solution in the labor market is an interior solution for the sorting problem of the self-employed. Given finite wages, the sorting problem of potential entrepreneurs also has an interior solution.

(b) *Consumers' policy and value functions are linear in wealth and therefore aggregate. There is a single stochastic discount factor* $S_{t,t+1}$ *in the economy.*

(c) *Aggregate production function: The production functions of entrepreneurs aggregate to an aggregate production function with endogenous productivity:*

$$Y_{e,t} = AZ_{e,t}^{1-\phi}\big(K_t^{\alpha}N_t^{1-\alpha}\big)^{\phi}$$

Productivity $Z_{e,t}$ *follows an AR(1) process given by:*

$$Z_{e,t} = (1-\theta)g_{e,t}Z_{e,t-1} + Z_{e,0,t}$$

where $Z_{e,0,t} = \epsilon\theta e^{\mu_t}\int_{\bar{O}_{e,t}}^{\infty}\int_{\bar{U}_t(\tilde{O})}^{\infty}\tilde{U}\tilde{O}dF_U(\tilde{U})dF_O(\tilde{O})$*. The effective measure of potential entrepreneurs in the labor market* N_t^x *also follows an AR(1) process:*

$$N_t^x = (1-\theta)N_{t-1}^x + \epsilon\theta\big(F_{O,e}(\bar{O}_{e,t}) + (1-F_{O,e}(\bar{O}_{e,t}))EX_t(1-\eta)\big).$$

Proof: see Appendix A.1

Appendix B.8 lists all equilibrium conditions. It shows that only three endogenous state variables are needed to characterize the Markov equilibrium of this economy: past capital to characterize the capital goods sector, past labor supply, and past productivity of entrepreneurs to characterize the entrepreneurial sector.

The cutoff property of the solution is intuitive: Agents with higher observed skill O always do better either being entrepreneurs or self-employed than less skilled agents. In an interior equilibrium, there is therefore a cutoff above which all agents enter, $\bar{O}_t$. For entrepreneurs there is a also a cutoff level of risky productivity below which they exit, called $\bar{U}_t(O)$.

Proposition 1 allows me to determine prices in this economy for any given allocation. Part (b) of the proposition rests on the homogeneity of preferences and complete markets. The substantive point of part (b) is the existence of a common stochastic discount factor $S_{t,t+s}$. Given $S_{t,t+s}$ and the state of the economy, the rental rate of capital follows from a standard Euler equation. The rental rate of capital and the discount factor together with the wage allow me to characterize the entrepreneurial sector.

The cutoff property of the occupational choice for the self-employed pins down the wage. The entry threshold is such that the marginal self-employed is just indifferent between being a worker or self-employed for a given wage W_t^f:

$$\bar{O}_{s,t} = \frac{W_t^f}{\omega_t} \qquad \Leftrightarrow \qquad W_t^f = \bar{O}_{s,t}\omega_t. \tag{3.18}$$

Labor supply is $F_O(\bar{O}_{s,t}) = F_O\big(\frac{W_t}{\omega_t}\big)$. When the wage rate is high, more self-employed give up their business and become workers.

The crucial feature for the aggregation of entrepreneurial production is that entrepreneurs can hire all inputs on competitive spot markets, after observing their productivity. State by state, their factor demands are linear in their productivity. Factor market clearing then allows me to re-write the production function as in part (c) of the proposition. My model is therefore a model of intangible capital, as in Hall (2001). Hall's intangible capital is the endogenous total productivity of entrepreneurs in this model.

The simple equation for the evolution of total productivity $Z_{e,t}$ follows from *iid* exit and growth of firms. A law of large numbers then implies the evolution for total productivity as a first order autoregressive process. The productivity parameter $Z_{e,0,t}$ of entering entrepreneurs is the innovation to this AR(1) process:

$$Z_{e,0,t} = \epsilon\theta e^{\mu_t} \int_{\bar{O}_{e,t}}^{\infty} \int_{\bar{U}_t(\tilde{O})}^{\infty} \tilde{U}\tilde{O} dF_{U,t}(\tilde{U}) dF_O(\tilde{O}) \tag{3.19}$$

Z_{e0t} depends on exogenous risk via $F_{U,t}$ and endogenous cutoffs. The productivity distribution is given in closed form (A.3) and derived in Appendix B.3. The remainder of this section focuses on the determinants of the cutoffs.

Only the entry cutoff adjusts to macroeconomic conditions other than risk. In contrast, the marginal exit cutoff and the exit rate only depend on risk r_t, ℓ_t. For entrepreneurs, a sufficient statistic for macroeconomic conditions is the relative present discounted value of wages relative to profits.

Lemma 1. *In an interior equilibrium, the following properties hold:*

(a) The cutoff for entrepreneurial entry $\bar{O}_{e,t}$ is a function only of risk r_t, ℓ_t and macroeconomic conditions summarized in $\pi_t e^{\mu_t}$, where $\pi_t \equiv \frac{PDV_t(G_e\Pi)}{PDV_t(W)}$, where $PDV_t(\cdot)$ denotes the expected discounted value of a payment stream.[15] *In particular:*

$$\bar{O}_{e,t} = \frac{g(r_t, \ell_t)}{\pi_t e^{\mu_t}}. \tag{3.20}$$

(b) The exit rate EX_t is independent of macroeconomic conditions other than risk if $\bar{O}_{e,t} > 1$:

$$EX_t = \chi(r_t, \ell_t). \tag{3.21}$$

Proof: Appendix A.2.

[15] Given the stochastic discount factor $S_{t,t+s}$: $PDV_t(W) = \mathbb{E}[\sum_{s=t}^{\infty} S_{t,s} W_s | \mathcal{F}_t^-]$. For profits, define $G_{i,e,0\to t} \equiv \prod_{s=0}^{t} G_{i,e,s}$ as the cumulative individual productivity growth. Then $PDV_t(G_e\Pi) = \mathbb{E}[\sum_{s=t}^{\infty} S_{t,s} G_{i,e,0\to t+s} \Pi_{t+s} | \mathcal{F}_t^-] = \mathbb{E}[\sum_{s=t}^{\infty} S_{t,s} \bar{g}_e^{t+s} \Pi_{t+s} | \mathcal{F}_t^-]$ because $G_{i,e,\circ}$ is *iid* with mean g_e.

Note that part (a) of Lemma 1 is independent of the particular distributional assumptions: It holds more generally for any distribution $F_U(u|\theta)$ guaranteeing the existence of the risk-adjusted expectation in (3.23), with $g(r, \ell)$ replaced by $\tilde{g}(\theta)$. Part (b) relies on the specific distributional assumptions, but finding that exit is independent of macroeconomic conditions other than risk is consistent with evidence from the manufacturing industry that exit rates are approximately acyclical (Lee and Mukoyama, 2008).

The main steps in the proof of Lemma 1 highlight the model mechanics. Start with part (a). Potential entrepreneurs enter as long as their risk-adjusted lifetime value is higher as entrepreneurs than as workers. I show in Appendix A.2 that this choice is equivalent to choosing the occupation with the highest risk-adjusted wealth. For the marginal entrepreneur with skill $\bar{O}_{e,t}$, risk-adjusted wealth is therefore equal:

$$PDV_t(W) = \mathcal{R}[\max\{(1-\eta)PDV_t(W), \bar{O}_{e,t}e^{\mu_t}PDV_t(G_e\bar{\Pi})U_{i0t}\}|\mathcal{F}_{i0t}^{-}]. \tag{3.22}$$

Note that the risk-adjusted expectations operator $\mathcal{R}$ is homogeneous of degree one in all $\mathcal{F}_{i0t}^{-}$-measurable variables, and in particular, in the present discounted value of wages. I can therefore rewrite the indifference condition in terms of the relative profit ratio $\pi_t e^{\mu_t}$:

$$1 = \mathcal{R}\left[\max\left\{(1-\eta), \bar{O}_{e,t}\pi_t e^{\mu_t}U_{i0t}\right\}\middle|\mathcal{F}_{i0t}^{-}\right]. \tag{3.23}$$

Equation (3.23) shows that the relative profit ratio $\pi_t e^{\mu_t}$ is indeed a sufficient statistic for the macroeconomic state from the perspective of entrepreneurs. Note that only U_{i0t} is unknown under information $\mathcal{F}_{i0t}^{-}$. The distribution of U_{i0t} is constant when risk r_t, ℓ_t is constant. The entry threshold $\bar{O}_{e,t}$ adjusts therefore one-to-one to all anticipated changes in the relative profitability of entrepreneurs $\pi_t e^{\mu_t}$. Since the risk-adjusted expectation in (3.23) depends on (r_t, ℓ_t) and constant parameters only, equation (3.20) follows.

The left panel in Figure 4 illustrates how the entry cutoff varies with risk: It falls in upside risk and increases with downside risk for a given relative profit ratio $\pi_t e^{\mu_t}$.

Unlike the result on the entry threshold, part (b) of Lemma 1 exploits that the skill distribution conditional on entry follows a Power Law, given $\bar{O}_{e,t} > 1$. As macroeconomic conditions $\pi_t e^{\mu_t}$ improve, it then follows that entrepreneurial skills scale down proportionally from equation (3.20), while their rescaled distribution is unchanged. All else equal, this selection effect would increase the exit rate. However, improved macroeconomic conditions exactly offset the effect of proportionally lower skills on the exit threshold. Since the rescaled distribution and the new exit threshold are unchanged, so is the exit rate.

To see the offsetting effect on the exit rate, note that entrepreneurs exit whenever

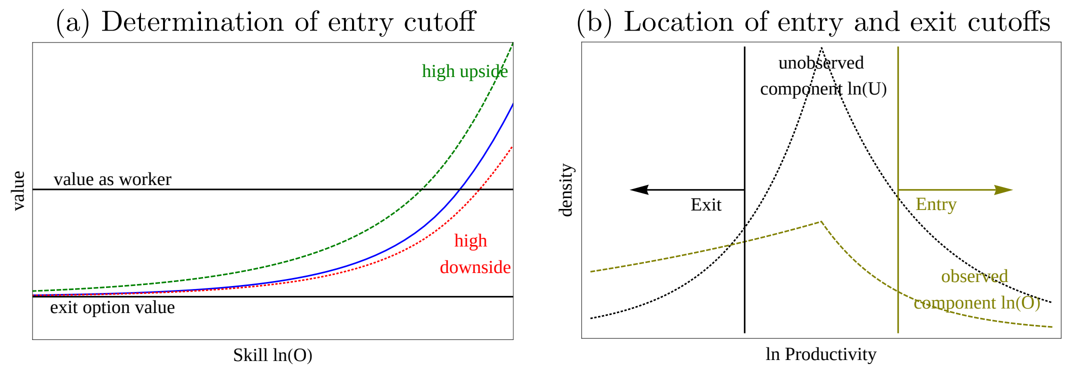


Figure 4: Determination of entry cutoff (left) and location of entry and exit cutoffs
Panel (a) illustrates how the entry cutoff depends on risk: It increases with upside risk r_t^{-1} and falls with downside risk ℓ_t^{-1} for a given relative profit ratio $\pi_t e^{\mu_t}$ (i.e., in partial equilibrium). Potential entrepreneurs compare the risk-adjusted value of entry with their outside option as workers. The outside option is independent of their skill. The value of entry is bounded below by the value of the exit option. Higher upside risk (green dashed line) raises the value of entry uniformly compared to the baseline (blue solid line). The marginal entrepreneur has a risk-adjusted value equal to the value of being a worker. Panel (b) illustrates the assumption on entry and exit cutoffs: Few potential entrepreneurs enter, so that entry is to the right of the mode of the observed skill distribution. Similarly, exit is relatively uncommon.

their productivity draw is too low: $U_{i0t} \leq \frac{(1-\eta)}{O_{i0t}\pi_t e^{\mu_t}} \quad \forall O_{i0t} \geq \bar{O}_{e,t} > 1$. Since skills O_{i0t} are scaled down by $\frac{1}{\pi_t e^{\mu_t}}$, the effective exit cutoff is unchanged. For example, the highest exit cutoff is associated with the lowest skill and satisfies:

$$\bar{U}_t\left(\bar{O}_{e,t}\right) = \frac{(1-\eta)}{\bar{O}_{e,t}\pi_t e^{\mu_t}} = \frac{(1-\eta)}{g(r_t, \ell_t)}. \tag{3.24}$$

As skills are proportionately higher, the cutoff is proportionately lower, but it is still independent of macroeconomic conditions other than entrepreneurial risk r_t, ℓ_t. This result highlights how risk is identified from data on exit rates in the model estimation.[16]

In the remainder of this paper, I assume that the equilibrium is in a specific region of the parameter space where entry and exit are relatively uncommon events: The modal potential entrepreneur shies away from entry, the modal entrant does not exit. While neither assumption seems restrictive for observed entry and exit rates, the assumption on entry is important for the exit characterization here. The second assumption only simplifies the algebra while maintaining the qualitative equilibrium properties.[17]

[16]Note from the proof that while entrepreneurial risk r_t, ℓ_t as a state variable typically affects π_t through expected future entry and exit, the influence on the exit rate and the entry cutoff is only contemporaneous via its influence on the cross-sectional distribution.

[17]With higher risk aversion, the relative risk-adjusted benefit of entrepreneurship falls below one. Hence, part (a) of the assumption can be satisfied when agents are sufficiently risk-averse. By increasing $\bar{O}_{e,t}$, risk aversion also lowers the exit cost required to satisfy (b). In general, (b) is satisfied as $\eta \nearrow 1$.

Assumption 3. *The exit cost η and the risk aversion ρ are high enough to guarantee the following conditions:*

(a) The log entry cutoff is strictly positive: $\ln \bar{O}_{e,t} > 0$.

(b) The log exit cutoff is strictly negative: $\ln \bar{U}_t(\bar{O}_{e,t}) < 0$

3.4 Balanced growth

A balanced growth path exists in the economy, if productivity of the self-employed and entrepreneurs grows at the appropriate rates. Along the balanced growth path, aggregate labor demand is constant, while the aggregate demand for capital grows and the capital-to-labor ratio is growing. The capital share in wealth is constant.

Recall that the productivity growth of self-employed and entrepreneurs is $g_{s,t}$ and $g_{e,t}$, respectively. If Assumption 4 is satisfied, the share of the self-employed and entrepreneurs is constant along the balanced growth path, while physical output and its components as well as overall wealth grow at the common rate $g_{s,t}$.[18]

Assumption 4. *The expected growth of firms $\bar{g}_e$ equals the growth of household productivity $\bar{g}_s$, adjusted for decreasing returns to scale in the firm sector:* $\bar{g}_s = \bar{g}_e^{\frac{1-\phi}{1-\alpha\phi}}$.

3.5 Productivity and employment size distribution

In equilibrium, the productivity distribution of entrants differs from the distribution of potential productivity because of the self-selection of entrepreneurs. Figure 5 plots the equilibrium density of log productivity $\ln Z$ in the cross-section. There are two selection effects. The selection of entrepreneurs into entry implies that the density of observed productivity $\ln Z$ shifts to the right. Graphically, this corresponds to the difference between the black dotted density of potential productivity and the solid blue density of observed productivity $\ln Z$ without exit. The second selection effect truncates the observed productivity distribution at the minimum efficient operating level. This effect is visible in the difference between the green dashed and blue solid curves.

The productivity distribution of entrants depends on macroeconomic conditions besides risk. When the relative profit ratio $\pi_t e^{\mu_t}$ is high, less skilled entrepreneurs also enter and less productive entrepreneurs continue operating. Looking at the (untruncated) den-

[18] This assumption is unnecessarily restrictive: The only necessary condition for asymptotically balanced growth is that the long-run trend of home and firm productivity are the same. The more restrictive assumption guarantees that the economy is on the balanced growth path absent other shocks.

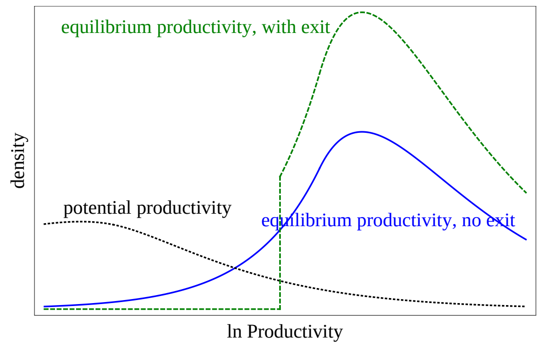


Figure 5: Equilibrium cross-sectional density of log productivity among entrants
Equilibrium productivity differs from potential productivity due to self-selection into entrepreneurship and exit. The right tail of the equilibrium productivity distribution is Pareto distributed. The dotted line is the density of potential productivity, the blue solid line the density of $\ln Z$ after selection into entry, but without an exit option. The green dashed lines adds the exit option to the blue line.

sity of log productivity shows that the distribution depends on the endogenous relative profit ratio π_t, besides the exogenous risk r_t, and ℓ_t:

$$f_{\ln Z,t}(z) = \frac{r_t\ell_t}{r_t+\ell_t}\frac{\nu}{\nu-r_t}\begin{cases} e^{\ell_t\left(z-\ln\frac{g(r_t,\ell_t)}{\pi_t}\right)}\frac{\nu-r_t}{\nu+\ell_t} & e^z\le \frac{g(r_t,\ell_t)}{\pi_t},\\ e^{-r_t(z-\ln\frac{g(r_t,\ell_t)}{\pi_t})}\left(1-\frac{\ell_t+r_t}{\nu+\ell_t}e^{-(\nu-r_t)(z-\ln\frac{g(r_t,\ell_t)}{\pi_t})}\right) & e^z>\frac{g(r_t,\ell_t)}{\pi_t}.\end{cases} \tag{3.25}$$

A higher profit ratio π_t shifts the (log) productivity distribution to the left, all else equal. Changes in μ_t, in contrast, have no direct effects on the distribution on entrants: The location parameter μ_t cancels out in (A.3) and affects only entry directly.

The employment size distribution among firms in any given state is equal to the cross-sectional distribution of productivity: Size is proportional to productivity of firms within a period (see Appendix A.1). The same is not true when comparing the size distribution over time. Over time, changes in the state of the economy also change how many workers an entrepreneur with given productivity hires, blurring the link between productivity and size. I therefore consider a special case with a quasianalytical solution to highlight the main driving forces before discussing the numerical solution for the general case.

3.6 Analytical solution: labor-only

This section derives quasianalytical solutions for macro aggregates under the simplifying assumptions of symmetric growth-risk exposure for workers and entrepreneurs and equal productivity of all self-employed. I use symmetric growth risk to describe the case of wages and profits growing at a common, possible stochastic rate. In this case, partial equilibrium analytical results carry over to general equilibrium. The closed economy with endogenous labor supply behaves as a small open economy with migration of workers.

Assuming symmetric growth-risk exposure makes the forward-looking entrepreneurial decision static and greatly simplifies the analysis. The economic content of this assumption is that the idiosyncratic risk at the start of a new business dominates the exposure to aggregate risk and uncertain growth later on in the firm's lifecycle. An alternative is to consider the more general case, but to solve the model numerically. This is done in the empirical section, but the numerical approximation procedure utilizes the same conceptual idea: Idiosyncratic risk matters most to entrepreneurs. Formally, I am imposing the following additional assumption:[19]

Assumption 5. *(a) Nonentrepreneurs have the same productivity as the self-employed:* $O_{i0} = 1$ *for all potential self-employed.*

(b) Labor is the only input into entrepreneurial production: $\alpha = 0, \delta_t = 1$.

(c) The economy is on the balanced growth path (Assumption 4).

(d) There is no wage rigidity: $\kappa = 0$.

(e) There is an interior equilibrium in the labor market: $Z_t^e \bar{N}_t \in (N_t^{\chi}, 1) \forall t$.

The forward-looking occupational choice for potential entrepreneurs becomes static because profits and wages grow at the same stochastic trend. Their relative present discounted value then equals the static relative profit ratio: $\pi_t = \frac{PDV_t(W)}{PDV_t(G_e\bar{\Pi})e^{\mu_t}} = \frac{W_t}{\Pi_t e^{\mu_t}}$.

Corollary 1. *Under Assumption 5, the equilibrium is as follows: The wage is given by the outside option:*

$$W_t = \omega_t. \tag{3.26}$$

The entry threshold for potential entrepreneurs satisfies:

$$\mathcal{R}[\max\{(1-\eta), \bar{O}_{e,t} U_{i0t} \frac{\bar{\Pi}}{\bar{\omega} e^{\mu_t}}\} | \mathcal{F}_{i0t}^-] = 1, \tag{3.27}$$

where $\bar{\omega}$ *is the detrended wage rate and* $\bar{\Pi}$ *satisfies from Appendix A.1:*

$$\bar{\Pi} = (1-\phi)\phi^{\frac{\phi}{1-\phi}} \bar{\omega}^{-\frac{\phi}{1-\phi}}. \tag{3.28}$$

[19] Relaxing parts (a) through (d) of Assumption 5 leaves the character of the solution unchanged but requires a numerical solution. Assumption (e) implies that the scarce factor in the economy is entrepreneurial talent, whereas labor supply is relatively abundant. If Assumption (e) were violated, the character of the solution would change fundamentally. Either no one would be an entrepreneur or labor would be rationed with wages rather than quantities adjusting to clear the labor market.

The exit threshold is given by:

$$\bar{U}_t(\bar{O}_{e,t}) = (1-\eta)\frac{\bar{\omega}}{\bar{\Pi}e^{\mu_t}\bar{O}_{e,t}} \tag{3.29}$$

Proof: See below.

The corollary follows from Proposition 1 after noting that with homogeneous workers an interior equilibrium requires the wage to equal the common outside option from (3.18). Given balanced growth, the sum of profits of entrepreneurs and workers grow at the same rate. Detrending equations (3.23) and (3.24) yields equations (3.27) and (3.29), respectively.

Because the cutoff is static, aggregate equilibrium allocations are not forward-looking. Expectations of the future matter only for individual allocations and asset prices. Agents demand a risk premium for aggregate risk. Young entrepreneurs trade claims with other agents to smooth consumption.

Table 1 provides an analytical characterization of the dynamic impacts of changing risk in this simplified economy. The lower part of the table provides additional comparative static results. I distinguish static effects from dynamic impacts since in the equilibrium considered here also the impact response to changes in the productivity distribution can be computed analytically, without taking a stand on the specific properties of the underlying stochastic process.

The logic behind the results in Table 1 is straightforward. Consider the effect of upside

Table 1: Effects of risk changes for interior, labor only equilibria

	Increase in ...	Entry (1)	Average size (2)	Job creation (1) × (2)
Dynamic impact	Upside risk r^{-1}	$+$	$+^a$	$+$
	Downside risk ℓ^{-1}	$-$	$-^a$	$-$
	Overall risk	$-/+^b$	$-/+^b$	$-/+^b$
	Relative upside risk	$+$	$+/-$	$+$
	Multiplicative mean $\bar{\mu}$	$+$	0	$+$
Static effect	Risk aversion ρ	$-$	$+^c$	$-$
	Outside option (wage) $\bar{\omega}$	$-$	0	$-$

A more favorable risk composition increases entry and overall productivity because of first order stochastic dominance shifts. Deterministic shifts in the productivity distribution or the outside option leave average size unchanged. If risk aversion dominates the exit option, average size increases when the risk composition becomes more favorable or the agent is more risk averse. Here, overall risk is defined as $\omega = \sqrt{\ell_t^{-1} r_t^{-1}}$. Relative risk $\upsilon = \sqrt{\ell_t / r_t}$. [a] sufficient condition: $\eta \approx 1$. [b] sufficient condition for positive effects: $\eta \approx 1$ and $1+\upsilon^2-\omega\upsilon(\upsilon^2-1)(\rho-2)+2\omega^2\upsilon^2(\rho-1)>0$. [c] sufficient condition: $\ell+r+\nu-r\nu>0$. Proof: Appendix A.4

risk. An increase in upside risk makes the distribution of productivity more favorable and causes lower skilled entrepreneurs to enter. Overall entry rises. However, since agents are risk averse, the entering types' productivity is not low enough to offset an increase in average productivity. For both reasons, overall productivity and therefore employment increases. The effect of an increase in downside risk is qualitatively symmetric.

To analyze what happens if both tail risks increase at the same time I introduce the concept of overall risk versus relative risk. I define overall risk as the geometric average of the two tail coefficients: $\sqrt{\ell_t^{-1} r_t^{-1}}$. By analogy, I define relative upside risk as $\frac{\sqrt{\ell_t}}{\sqrt{r_t}}$. The effect of an increase in overall risk is ambiguous and depends on whether relative upside risk is high. If risk aversion is high enough, or relative upside risk is low enough, entry and overall productivity falls, while average productivity can increase.

Higher risk aversion or a lower outside option imply a change in the stationary equilibrium toward a smaller entrepreneurial sector. When risk aversion is higher, the requires risk premium increases. Since the risk premium equals $\mathbb{E}\left[\max\{(1-\eta), \bar{O}_{e,t} U e^{mu_t} \pi_t\} | \mathcal{F}_t\right] - 1$, it is immediate that a higher risk premium is equivalent to a larger expected average firm size for the marginal entrepreneur. Consequently, average firm size increases. Higher wages increase the average productivity but leave the average employment size unchanged.

Three features of the model are key to these results: Exit costs, risk aversion, and first order stochastic dominance (FOSD) shifts. Lemma 5 in the appendix proves that increases in upside risk cause a FOSD shift in the risky productivity distribution and vice versa for downside risk. If the productivity distribution improves in a FOSD sense, overall productivity would increase given entry. However, since entrepreneurship becomes more attractive relative to being a worker, entry also increases and lowers the immediate effect of a FOSD on the average productivity. Indeed, the average productivity remains unchanged with a deterministic FOSD shift (an increase in the location parameter μ_t): The entry of lower skilled entrepreneurs offsets the increased average productivity of supra-marginal entrepreneurs. Only risk aversion or exit drive a wedge between average productivity and the payoff to an entrepreneur, causing changes in average productivity and average size.[20]

[20] For the mathematical intuition without exit, note that average productivity equals $\mathbb{E}[Z_{i0t}] = \frac{\bar{O}_{e,t}}{1-\nu^{-1}} \mathbb{E}[U_t]$, given the conditional Pareto distribution of skills. In equilibrium, the skill cutoff adjusts: $\mathbb{E}[Z_{i0t}] = \frac{\mathbb{E}[U_t]\bar{\omega}}{(1-1-\nu^{-1})\mathcal{R}[U_t] e^{\mu_t} \bar{\Pi}}$ and is independent of μ_t. While average productivity falls with an increase in $\bar{\omega}$, profits $\bar{\Pi}$ and unit labor demand $\bar{N}$ fall also to exactly offset the effect on average employment size. See Appendix A.4 for the general case.

4 Application to US data

I use data on the size distribution of the universe of US businesses from 1976 to 2011, as provided in the Business Dynamics Statistics (BDS). The BDS data is available both at the establishment and at the firm-level. The data is described in Appendix C.6.

4.1 Measuring tail risk

I prove that data on the observed size distribution and exit rates allows me to infer ex ante upside and downside risk faced by entrepreneurs in my model. While the identification of downside risk depends on specific parametric assumptions, I also discuss how the inference of upside risk generalizes.

My identifying assumption to recover upside risk is that the difference in size among the largest entrants is primarily driven by risky productivity draws. Entrepreneurial skills also matter, but they are second order when conditioning on the largest businesses only: The heterogeneity in skills among the most successful entrepreneurs is dominated by the heterogeneity of their productivity draws. Formally, this means that the tail coefficient of productivity risk is larger than that of the skill distribution: $r_t^{-1} > \nu^{-1}$.[21]

Upside risk is the fatness of the entrants' size distribution. To fix ideas, assume for now that there are no skill differences among entrepreneurs. Then the cross-sectional productivity distribution in my model would be exactly Pareto above its mode z_{mode} and its survivor function an exact power law:

$$\bar{F}_{Z|Z>z_{mode}}(z) = 1 - F_{Z|Z>z_{mode}}(z) = \left(\frac{z}{z_{mode}}\right)^{-r_t}. \tag{4.1}$$

Because all measures of firm size in my model are proportional to productivity, the Pareto property carries over to employment size.

The Pareto distribution is a good first order approximation to the upper tail of the size distribution of new businesses. If the Pareto distribution held exactly in the data, equation (4.1) would imply that the log empirical survivor function and the log of business size should be linearly related to one another. As the left panel in Figure 6 shows, this relationship holds well in the BDS data for new establishments. The appendix plots the corresponding data for more years (Figure C.11). The associated R^2 is about 0.99 in

[21] To assess this assumption, it would be desirable to compute the tail index conditional on entrepreneurial characteristics such as demographics, or finances. But the Survey of Consumer Finances, for example, includes fewer than 100 nonimputed businesses in each wave that are less than three years old, making sample splits in the tails infeasible.

every year from 1977 to 2009.[22] The right panel shows the analogue for data on new firms. Here, the distribution of the largest firms curves inward for the extreme tail of the distribution. This pattern is not unusual for the tail of the distribution where the density of the distribution is low.[23]

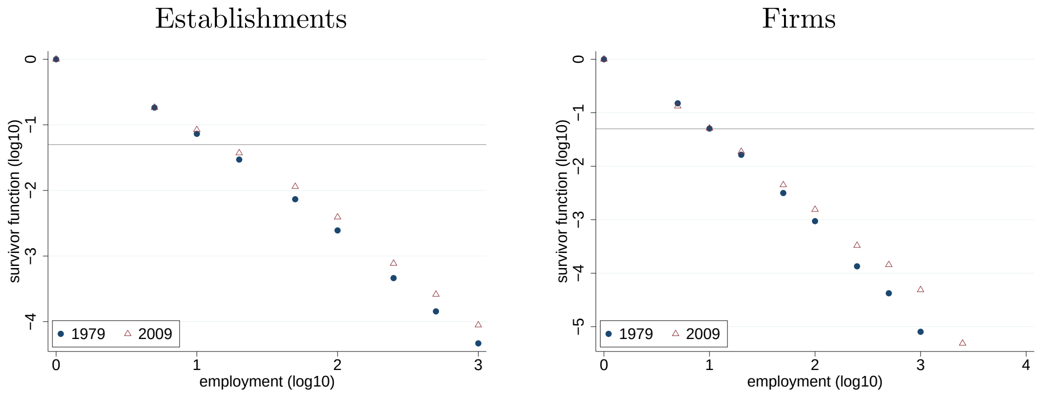


Figure 6: Log counter-cumulative distribution function vs. log size for businesses age 0
A Pareto distribution is a good first order approximation to entrants' size distribution. Differences in slope are evidence for time variation in upside risk. Plotted is $\log_{10} \bar{F}(z)$ vs. $\log_{10} z$, where z represents employment.

While the Pareto distribution is a good first order approximation, it neither holds exactly in my model nor in the data. Particularly, my model implies that the survivor function for productivity has the following form:

$$\bar{F}_{Z,t}(z) = z^{-r_t}(\kappa_{1,t} + \kappa_{2,t} z^{-(\nu - r_t)} + o(z^{-(\nu - r_t)})), \kappa_{1,t} > 0 \forall t. \tag{4.2}$$

For high productivity firms, $\bar{F}_{Z,t}$ differs from an exact Pareto survivor function by an approximately constant function only. It therefore belongs to the so-called Hall (1982) class of distribution function. In my data, I only observe the empirical distribution function of entering firms. Welsh (1986) shows how to estimate the tail coefficient r_t^{-1} from the empirical distribution function using a simple slope estimator. My model therefore allows me to recover the right tail coefficient from the reported data on entrants' size

[22] The fit for 2010 drops to an R^2 of only 0.8. This may be because of delays in capturing or classifying establishment-level data at the end of the sample.

[23] The literature considers the firm size distribution across all ages. Luttmer (2007, Figure 1) for 2000 and Rossi-Hansberg and Wright (2007, Figure 1) for 2002 coincide in the good fit of a power law distribution for firms up to about 10,000 employees. However, Rossi-Hansberg and Wright (2007) plot the distribution above the range of 100,000 employees where the pattern deviates from a power law similar to the pattern in the right panel of Figure 6. Note that Rossi-Hansberg and Wright (2007) argue that establishments are better described by a log-normal distribution when averaged across all ages, while Oberfield (2011) finds a heavy tail for manufacturing establishments. I find that for young establishments a power law holds well over the relevant range of the distribution.

distribution. Lemma 1(b) then implies that downside risk can be backed out from the estimate r_t^{-1} and the observed exit rate. Proposition 2 formalizes this statement.

Proposition 2. *Assume that $r_t^{-1} > \nu^{-1}$. The model in Section 3 then implies that the following estimator for upside risk is consistent and asymptotically normally distributed. Fix $u > 1$. If $z \propto m^{-1} n^{\frac{1}{2\nu-\omega}}$, then as $n \to \infty$:*

$$n^{\beta}\left(\frac{\ln(\bar{F}_n(uz)) - \ln(\bar{F}_n(z))}{-\ln(u)} - r_t\right) \xrightarrow{d} \frac{m^{-\frac{r_t}{2}}}{\ln(u)}\mathcal{N}\left(0, \frac{1-u^{-r_t}}{\kappa_{1,t}}\right) + \frac{\kappa_{2,t}}{\kappa_{1,t}}\frac{m^{\nu-r_t}}{\ln(u)}(u^{\nu-r_t}-1), \tag{4.3}$$

where $\kappa_{i,t} = \kappa_i \bar{N}_t^{-1}, i = 1,2$ and $\beta = \frac{\nu - r_t}{2\nu - r_t} \in (0, 0.5)$.

Given exit rates and estimated upside risk, downside risk can be backed out from the exit rate according to equation (3.24): $EX_t = \chi(r_t, \ell_t)$.

Proof: Appendix A.3.

Being asymptotic in its nature, the Proposition is silent about the small sample properties of the estimator. I therefore follow the literature, e.g. Kelly (2012), in considering the slope in the upper 5% or 1% of the sample. This means measuring the slope between the empirical frequency of businesses with 20+ employees relative to those with 50+ employees. This cutoff always lies between 2.9% and 4.6% of the sample of new businesses. I also consider the slope measured between businesses with 50+ and 100+ employees. These ranges seem relevant for the most successful entrepreneurs: The benchmark start-up in the "Doing Business" report (World Bank, 2013) has 10 to 50 employees (p. 109). I also consider a GLS-type estimator for upside risk that uses the entire upper tail for robustness. This estimator weighs the data points with the inverse of their asymptotic variance, corresponding to the number of businesses in a given bin.

I choose a simple estimator such as the one proposed by Welsh (1986) over full maximum likelihood estimation of my model, because the identification of upper risk in my model transcends the particular functional forms I have used. Tail risk is first order for entrepreneurs in my model, while small deviations from a Pareto-distributed upper tail risk can be expected to have only second order effects. Focusing on the tail behavior then allows consistent estimation for all distribution function in the Hall (1982) class obeying equation (4.2).[24]

Figure 7 plots the estimated time series for upside risk, using both establishment and firm-level data. It considers two cutoffs: 20 and 50 employees. When considering a cutoff of 20+ employees, most raw measures of upside risk exhibit a trend over the sample. This may reflect shifts either in the composition of new businesses from manufacturing toward

[24] This is similar to the argument in Hill (1975) who justifies conditioning on the tail by appealing to a diffuse prior over the specific unconditional distribution.

the service sector over time or the time-varying number of new businesses in the sample. Either can introduce a trend due to specification error. I therefore detrend the estimates of upside risk when necessary, but I focus on risk measures at a higher cutoff without detrending to ensure that my results are not driven by the trend removal.

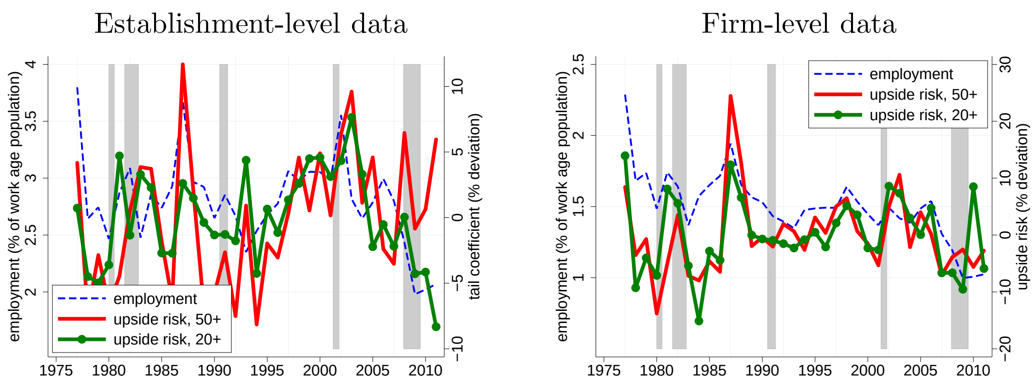


Figure 7: Upside risk and employment at businesses age 0, 1977–2011

Estimates of upside risk $\ln \hat{r}_t^{-1}$ exhibit significant time variation, with highs in 1977, the mid 1980s, and the late 1990s. Raw correlations of employment with risk measures are 0.44 (50+) and 0.48 (20+) for establishment-level data and 0.49 for firm-level data for either cutoff. Upside risk is measured at two different cutoffs, 20 or 50 employees. Upside risk at 20+ employees is in deviations from linear trend. Shaded areas indicate NBER recession dates.

4.2 Structural estimation

To solve for the full model dynamics with capital, I log-linearize before I estimate the model. My historical decomposition shows that risk can account for 40-54% of the observed variation in employment of young businesses. Both risk components help to explain employment at young businesses. [25]

4.2.1 Data used in estimation

In the estimation, I use data on employment at businesses aged up to one year.[26] This is because I observe only exit between age zero and one. To match the aggregated cohort data for employment, I focus here on the upside risk estimated at the initial

[25] I use Dynare (Adjemian et al., 2011) for the computations. The posterior is simulated starting at the posterior mode, uses three Metropolis-Hastings chains of length 30,000 plus a burn-in phase of 7,500 draws and a scale parameter of 0.55, yielding an acceptance rate around 0.3 and ensuring convergence of the Markov-Chain according to the Brooks and Gelman (1998) second moment and interval-based criterion. I initialize the Kalman filter at the stationary variance-covariance matrix. To find the posterior mode, I a standard Newton-type optimizer. The convergence proved robust to different starting values.

[26] All employment and entry data are detrended by dividing by the working age population.

age 0 averaged over the current and past year.[27] Section 4.2.6 considers cases without averaging as robustness checks, when upside risk is estimated either for businesses aged one or both employment and current risk are measured at age zero.

I use estimates of upside risk, the exit rate and employment of businesses up to age one, as well as employment growth of older businesses for the estimation. Four shock processes correspond to these measurement equations: upside and downside risk shocks, as well as shocks to the mean productivity of new businesses and the real wage shock. Because Proposition 2 implies that the upside tail coefficient is estimated with error, I allow for *iid* measurement error in the observation equation for upside risk approximately equal to $\sqrt{2}$ times the minimum estimation error over the sample to account for specification error and averaging. This effectively filters out some high frequency movements as measurement error, while attributing persistent changes in estimated risk to actual risk changes. For the estimates, however, this filtering effect is small, because the estimated variance of the measurement error is roughly one quarter of the variance of the systematic component. Appendix C.6 plots the raw data.

Note that employment at young firms exhibits a downward trend relative to the working age population, while measured upside risk exhibits an upward trend (Figure C.9). In the baseline firm-level estimates, I remove the trend from upside risk and leave employment unchanged. The results using both firm-level and establishment-level data are virtually unchanged when trends are removed from all variables prior to estimation (Figure C.13 in the appendix).

4.2.2 Calibration and estimation

I estimate the shock processes and the risk aversion that governs the response to risk shocks. The remaining parameters are calibrated. Priors for all parameters are independent and standard.[28] Priors for the standard deviations of shocks follow inverse gamma distributions with diffuse priors. Autoregressive coefficients are restricted to lie in $(0, 1)$ and follow a priori a beta distribution. Downside risk is inverse gamma distributed. Similarly, risk aversion is inverse gamma distributed, up to a location shift to $(1, \infty)$. Means in observation equations follow a diffuse prior and account for unmodeled life-cycle differences in exit rates for businesses.

[27] In line with Haltiwanger et al. (2010), I use the average number of new businesses as the denominator for the exit rate: $EX_t = \frac{\text{exiting businsses}_{t,t+1}}{0.5(\text{operating businesses}_t + \text{operating businesses}_{t+1}}$. Entry is then defined as the denominator of the exit rate over the working age population. Upside risk is $\ln(r_t - 1) = \frac{1}{2}(\ln(\hat{r}_{t-1} - 1) + \ln(\hat{r}_t - 1))$.

[28] Formally, the priors are not truly independent. As in the literature, the structural restrictions of the model are implicitly part of the priors. This includes Assumption 3 on risk aversion. When sampling from the posterior and estimating in practice, the restrictions are satisfied along each sample path.

Table 2 lists the calibrated parameters. While several parameters are standard,[29] the parameters characterizing the entrepreneurial sector warrant some discussion. The death probability θ is 10% per year, so new businesses contribute 10% of employment in the stationary equilibrium. The share ϵ of potential entrepreneurs is normalized to 50% of the population. The share of entrants in the overall population is set to 0.3%, and the entry parameter λ is adjusted accordingly. The self-employed represent 10% of the population in the stationary equilibrium. In the data, the share of the self-employed varies between 7% and 10% over the sample period.[30] The scale of production $\bar{\mu}$ is set to scale employment accordingly. With the mean tail coefficient $\bar{r} = 1.5$ taken from the data, the entry elasticity is set to $\nu = 2$ to ensure that the conditions for measurement in Proposition 2 are satisfied.[31] The exit cost $\eta = 0.35$ corresponds to the foregone college wage premium, which is roughly 50% (Fortin, 2006). Based on the 2010 US median wage of \$40,000, the implied discounted cost of a new business amounts to about \$100,000.[32]

Table 3 lists the posterior mean and standard deviation for the estimated parameters for two different datasets. The posterior persistence of most shock processes is slightly higher than the prior persistence, with the persistence of risk shocks implying a half-life of shocks of about 1.5 years. For firm-level data, the estimated measurement error is only about half that of the structural innovation, while it is two-thirds that for establishment-level data. Risk aversion at the posterior mean ρ is 2.4 or 3.0, depending on the dataset, within the range of values considered in models of entrepreneurship. For example, Heaton and Lucas (2004) consider values of risk aversion between 0.5 and 5, while Hall and Woodward (2010) consider a value of $\rho = 2$ as their benchmark. Average estimated downside risk $\bar{\ell}$ is slightly lower than the calibrated average upside risk $\bar{r}^{-1}$, implying a slight right skew of the log-productivity risk distribution.

To assess the validity of my model, I compare the size-distribution of entrants in the stationary equilibrium to the average size distribution of new businesses in the data.

[29] The discount factor for financial assets, $\tilde{\beta}$ is set to 0.95%, so that the risk-free interest rate on financial assets is about 5% annually. The intertemporal elasticity of substitution is set to unity. The labor share in income is $(1-\alpha)(1-\phi) = 0.65$, the profit share is $1-\phi = 0.15$, the depreciation rate is $\delta = 0.08$, and the adjustment cost on capital is $\bar{\Psi} = 0.7^{-1}$, matching the estimated elasticity of investment with respect to Tobin's Q in Christiano et al. (2005). The real rigidity is taken from Blanchard and Galí (2007), adapted to annual frequency: $\kappa = 0.9^4 \approx 0.65$.

[30] I use the annual average for all industries relative to total nonfarm employment. The inclusion of the farm self-employed seems appropriate as the farm sector does represent an outside option for workers.

[31] Note that the implied average tail index for young firms here implies a thinner tail than estimates for the size distribution averaged across all ages. This implies that the tail becomes fatter over time. This is consistent with my model when the growth process is specified appropriately. For example, when growth rates are log-normally distributed a Poisson death rate implies a fat tail of the size distribution (Reed and Jorgensen, 2004), which could be closer to a unit tail index as in the literature on old firms.

[32] The benchmark start-up capital in the World Bank (2013) "Doing Business" report is 10 times the average annual per capita income in a given country (p. 109), which equals the expected life-time income here; $\eta = 0.35$ implies that roughly one-third of the start-up cost is sunk.

Table 2: Calibrated parameters

Parameters	Value	Target
Exogenous exit θ	0.10	10% employment share
Potential entrepreneurs ϵ	0.50	Normalization
Share of entrepreneurs	0.003	BDS data
Share of self-employed	0.10	$7.5 - 10\%$ in data
Assets to human capital income	4.3	NIPA sample mean
Avg upside risk $\bar{r}$	1.5, 1.6	BDS data (establishments, firms)
Entry elasticity ν	2	Consistency with proposition 2
Cost of exit η	0.35	$1 - (\text{college wage premium})^{-1}$
Adjustment cost $\bar{\Psi}$	0.7^{-1}	Lucas-Prescott specification, Christiano et al. (2005)
Labor share $(1-\alpha)(1-\phi)$	0.65	Cooley and Prescott (1995)
Depreciation rate $\bar{\delta}$	0.08	Cooley and Prescott (1995)
Profit share $1-\phi$	0.15	Gomme and Rupert (2004)
Labor supply elasticity	3	Prescott (2004)
Real rigidity κ	0.65	Blanchard and Galí (2007)
IES ψ^{-1}	1	Constant savings rate
Discount factor $\tilde{\beta}$	0.95	5% risk free rate

Table 3: Priors and posteriors for establishment-level and firm-level data

Shock process		Prior mean (sd)	Posterior mean (sd)	
			Establishments	Firms
Upside risk $\ln(r_t - 1)$	AR(1)	0.5 (0.2)	0.72 (0.12)	0.62 (0.13)
	sd	0.2 (2.0)	0.07 (0.01)	0.10 (0.01)
iid measurement error	sd	0.05 (0.005)	0.05 (0.004)	0.05 (0.004)
Downside risk $\ln \ell_t$	AR(1)	0.5 (0.2)	0.63 (0.13)	0.59 (0.13)
	sd	0.2 (2.0)	0.08 (0.01)	0.07 (0.01)
New firms' productivity μ_t	AR(1)	0.5 (0.2)	0.69 (0.17)	0.85 (0.07)
	sd	0.2 (2.0)	0.04 (0.01)	0.04 (0.01)
Wage shock $\ln \omega_t$	AR(1)	0.5 (0.2)	0.60 (0.14)	0.65 (0.13)
	sd	0.2 (2.0)	0.02 (0.003)	0.02 (0.003)
Average downside risk $\bar{\ell}$		1.0 (2.0)	2.36 (0.53)	2.35 (0.42)
Risk aversion ρ		2.0 (2.0)	2.63 (0.40)	3.62 (0.86)

Priors for AR(1) parameters are Beta; priors for all other parameters inverse gamma. Risk aversion is transformed to $\rho - 1.01$ prior to estimation.

Figure 8 compares percentiles of the empirical distribution to model-implied percentiles, normalized to coincide for the largest businesses.[33] Although I estimate the parameters of my model using only a simple slope estimator for businesses with 50+ versus 100+ employees and time-series data, my estimates (in circles) tracks the overall shape of the

[33] Absolute firm size in my model is not readily comparable to the data, hence the normalization for businesses with 1,000 employees. I normalize to match the extreme to highlight potential differences in the slope. Note that the data is discretized to the number of paid employees. I therefore round the model-implied minimum scale, which fits the minimum size of one employee in the data mechanically.

empirical distribution (in squares) well. This is not mechanical: comparing my model with a differently calibrated slope parameter of $\bar{r} = 1.3$ implies percentiles that deviate significantly from the empirical percentiles.

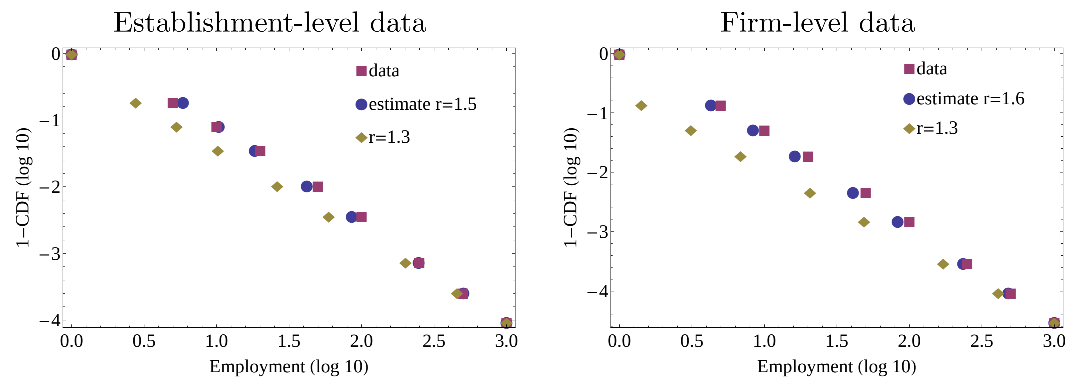


Figure 8: Comparison of model-implied and empirical stationary distribution of entrants
The model implied employment size percentiles (in circles) track the empirical percentiles (in squares) well. The model is chosen to fit largest percentile exactly and fits the minimum size mechanically when constraining the model to integers, in analogy to the data. Fitting the center of the distribution is not mechanical: Choosing a different tail coefficient of $\bar{r} = 1.3$ worsens the fit visibly.

4.2.3 Structural impulse-response functions

This section complements the analytical results obtained in section 3.6 by analyzing the response of the estimated model to risk shocks. While the basic logic of the analytical version without capital and heterogeneous workers carries through, the endogenous productivity dynamics lead to propagation and amplification of shocks. Risk aversion is important for the dynamic response to risk shocks.[34]

A one standard deviation shock in upside risk leads to the creation of 580,000 new jobs at new businesses on impact, about 0.2% of the working age population. It results in a gradual build-up of jobs in the overall economy, peaking only at about 280,000 new jobs after three years because of general equilibrium effects. This is shown in Figure 9, where the solid blue line shows employment at young businesses and the red crossed line indicates overall employment. The hump-shaped response of overall employment is typical for the response to persistent shocks in this economy. In the case of the one standard deviation shock to upside risk, the productivity of cohorts jumps up by 5% and declines only gradually over time. Initially, new cohorts are more productive than

[34]The results in this section are computed at the posterior mean. Appendix C.4 provides a complete set of impulse-response functions.

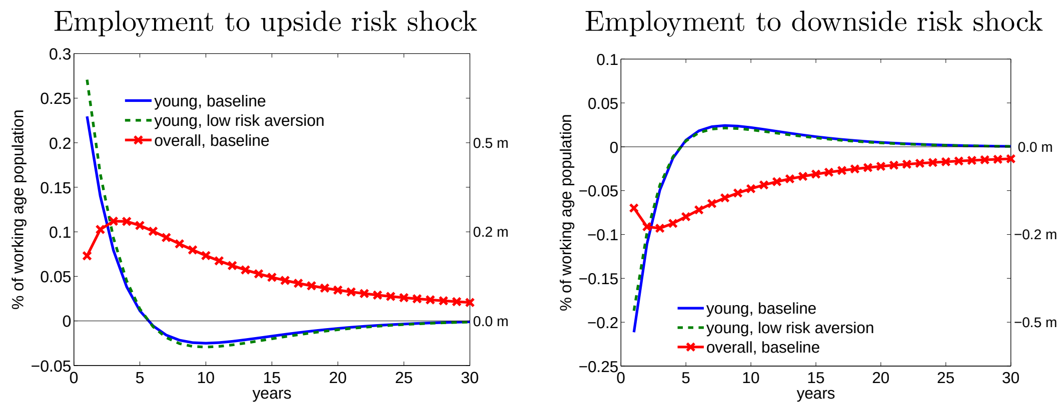


Figure 9: Impulse response of employment by young and all businesses to risk shocks
One standard deviation increases in upside (downside) risk cause increased (decreased) employment at young businesses. Risk aversion mutes the effect of upside risk shocks and amplifies the effect of downside risk. In general equilibrium, overall employment responds less as factor prices adjust. Baseline results evaluated at the posterior mean and the stationary equilibrium. Low risk aversion uses $\rho = 1$ instead of the posterior mean ρ. 0.1% of the working age population corresponded to 238,000 jobs in 2010.

the average exiting firm they replace. While this effect lasts, overall entrepreneurial productivity rises initially by 0.5% and peaks at 1% after four years (Appendix C.4).

The response of overall employment to risk shocks is sensitive to the specification of the labor market: If factor prices were fixed, overall employment would peak at roughly 1%, or 1.2 million jobs.[35] With flexible factor prices, however, the increased factor demand causes factor prices to increase. These increases reduce the peak impact from 1.2 million jobs to 280,000. The partial adjustment form of real wage rigidity considered here matters only in the short run: After three years, wages have mostly adjusted, but the impact response of overall employment to risk shocks would only be about half as high without the real wage rigidity (see Appendix C.4). Wage rigidities have no significant quantitative effect on employment by young businesses since productivity changes dominate for these businesses.

Risk aversion controls the magnitude of the response to risk shocks: Risk aversion amplifies the effects of downside risk shocks, but it mutes the response to upside risk shocks. Figure 9 illustrates this for employment by young businesses by comparing the solid to the dashed line. If entrepreneurs had a lower risk aversion of $\rho = 1$ rather than the estimated $\rho = 2.63$, the employment response to the risk shocks would be 12% to 18% more expansionary. Risk aversion affects job creation through its effect on entry.

[35] The determinants of aggregate employment changes in response to productivity changes can be seen from a first order approximation of the labor market. For example, when labor is the only input ($\alpha = 0$) the equilibrium employment effect is given by $\frac{d \ln N_t}{d \ln Z_t} = \frac{\epsilon(1-\phi)}{1-\kappa}\left(\frac{\epsilon(1-\phi)}{1-\kappa} + 1\right)^{-1}$. Stronger DRS (lower ϕ), higher labor supply elasticity ϵ, and higher wage rigidity κ all increase the impact effect.

Risk aversion has no significant effects on the response of employment to nonrisk shocks around the stationary equilibrium as shown in Appendix C.4.[36]

Because only old agents own capital and their productivity is predetermined, aggregate shocks typically affect the wealth distribution across cohorts. Shocks that raise the productivity of new entrepreneurs also raise the relative wealth of that cohort. The effect of a shock to the real wage lowers the relative wealth of the young because they hold no physical capital. Similarly, an increase in the depreciation of capital only hurts the old and therefore raises the relative wealth of the young.

4.2.4 Historical decomposition

What role did productivity risk play in explaining employment dynamics at young businesses? In this section, I use the log-linearized model and the Kalman smoother to decompose observed and model-implied time-series into their components. I find that risk-shocks can explain 38% to 48% of employment movements at young businesses.

To understand the counterfactual experiments in this section, note that the historical decomposition yields a log-linear decomposition of the time series in the model. The time series are decomposed exactly into the four shocks and initial conditions. This allows for simple counterfactual experiments: For example, I can compute the employment caused by upside risk alone by zeroing out the other three shocks and initial conditions.

Figure 10 compares observed employment at businesses up to age one to risk-implied employment, for both establishment-level and firm-level data. The historical employment growth is shown as the black solid line, while the counterfactual employment due to only risk shocks is the brown dashed line. The correlation between risk-implied and actual employment is high, 0.72 at the establishment and 0.76 at the firm-level. Given that the standard deviation of employment is lower than the observed standard deviation of employment, the fraction of the variance explained by risk is only 38% and 48%, respectively. Risk-implied employment can account for most of the increase in employment in the late 1990s using either dataset. This boom was more pronounced in the establishment-level data.[37] Only about one-third of the decline in employment after the 2007 recession is explained by changing risk, leaving room for factors such as the collateral channel emphasized by Schott (2013). Appendix C.5 discusses how debt my model

[36] Appendix C.4 also shows the effect of *iid* shocks to upside and downside risk accompanied by a change in μ_t which makes them mean-preserving. Employment falls on impact in response to both shocks because risk-aversion dominates the convexity coming from the exit option.

[37] This may be due to the definition of young businesses here, which excludes firms older than one year. For example, if Amazon and eBay, which were founded in 1994 and 1995, respectively, were responsible for the expansion, they would only show up in the firm-level data until one year after hiring their first paid employee. However, if they had multiple establishments, they would continue to be reflected as young businesses in the establishment-level data.

could accommodate external debt. Since the decline of risk-implied employment in my model predates the recession, this decline may reflect worse technological opportunities (Gordon, 2012).

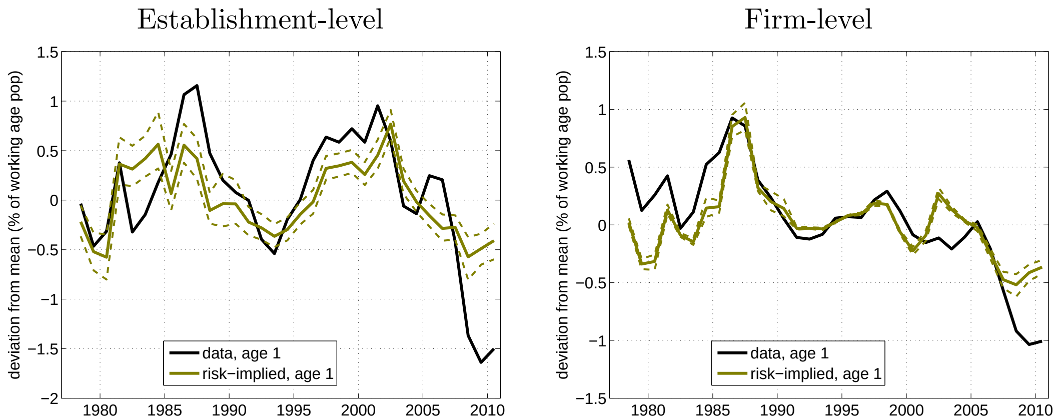


Figure 10: Actual and risk-implied historical employment up to age one, 1979–2011
Employment is measured at businesses up to age 1. Upside risk is the estimate for businesses aged zero, averaged over the current and past year. The cutoff for the risk estimate is 50 employees. Shown is the actual employment and counterfactual employment when non-risk shocks would have been absent.

Both risk components contribute to employment and both positive correlations and reasonable relative standard deviations account for the explanatory power. No component is excessively volatile. This is shown in Table 4, which shows the fraction of the historical variance explained by different shocks for different observables (first row for each observable) and the correlation (second row). Columns (1) and (2) list the contribution of upside and downside risk for different observables. Column (3) adds the two contributions up, whereas column (4) adds all shocks together. By construction, this adds up to 100% of the observed variance for observables used in the estimation. Shown are the posterior median and, in parentheses, the 10th and 90th percentile. Both upside and downside risk counterfactuals have significant and positive univariate correlation with observed employment, with its posterior median ranging from 0.38 to 0.63. Since the counterfactuals are less volatile than the data, they individually explain between 17% and 28% of the observed variation. Their joint explanatory power is approximately the sum of their individual contributions: Risk explains between 38% of employment of establishments and 48% for firms up to age one.

The table also contains additional information that can be used to assess the validity of the model, by comparing the model implied average size and entry to the data, which has not been used in the estimation. Here, the results for the full sample are mixed: At the establishment-level, risk implied entry and average size have a correlation of 0.57

Table 4: Historical decomposition of observed and model implied time series: 1979–2010

		(1) Upside	(2) Downside	(1) & (2)	All Shocks
		Establishment-level			
Employment,	rel var	0.17 (0.15,0.19)	0.22 (0.15,0.29)	0.38 (0.31,0.47)	1.00 (1.00,1.00)
age 1	corr	0.38 (0.33,0.44)	0.57 (0.54,0.62)	0.72 (0.65,0.76)	1.00 (1.00,1.00)
Employment	rel var	0.04 (0.03,0.05)	-0.01 (-0.01,0.00)	0.03 (0.02,0.04)	1.00 (1.00,1.00)
growth, age 2+	corr	0.20 (0.15,0.26)	-0.04 (-0.07,0.00)	0.14 (0.09,0.19)	1.00 (1.00,1.00)
Exit, age 1	rel var	0.04 (0.02,0.08)	0.96 (0.92,0.98)	1.00 (0.99,1.00)	1.00 (1.00,1.00)
	corr	0.36 (0.32,0.39)	0.99 (0.97,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.03 (0.01,0.04)	0.26 (0.17,0.38)	0.29 (0.19,0.41)	0.81 (0.72,0.91)
	corr	0.19 (0.07,0.34)	0.55 (0.54,0.56)	0.57 (0.54,0.60)	0.80 (0.76,0.83)
Average size	rel var	0.34 (0.30,0.38)	-0.02 (-0.04,0.00)	0.32 (0.28,0.36)	0.33 (0.23,0.45)
	corr	0.58 (0.53,0.61)	-0.19 (-0.38,0.05)	0.53 (0.45,0.59)	0.48 (0.33,0.59)
		Firm-level			
Employment,	rel var	0.28 (0.26,0.29)	0.20 (0.14,0.26)	0.48 (0.41,0.55)	1.00 (1.00,1.00)
age 1	corr	0.54 (0.51,0.58)	0.63 (0.60,0.68)	0.76 (0.71,0.79)	1.00 (1.00,1.00)
Employment	rel var	0.01 (0.00,0.01)	-0.05 (-0.06,-0.03)	-0.04 (-0.05,-0.03)	1.00 (1.00,1.00)
growth, age 2+	corr	0.02 (0.01,0.04)	-0.28 (-0.29,-0.27)	-0.12 (-0.14,-0.09)	1.00 (1.00,1.00)
Exit, age 1	rel var	-0.01 (-0.02,-0.00)	1.01 (1.00,1.02)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
	corr	-0.11 (-0.14,-0.08)	1.00 (0.98,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.06 (0.04,0.08)	0.25 (0.18,0.34)	0.31 (0.22,0.40)	0.90 (0.86,0.95)
	corr	0.38 (0.35,0.41)	0.56 (0.55,0.58)	0.64 (0.62,0.65)	0.89 (0.88,0.91)
Average size	rel var	0.25 (0.21,0.28)	-0.02 (-0.02,-0.01)	0.23 (0.20,0.26)	0.33 (0.30,0.37)
	corr	0.21 (0.20,0.23)	-0.07 (-0.10,-0.05)	0.21 (0.19,0.22)	0.27 (0.25,0.28)

For both firm-level and establishment-level data, both risk components of employment of young businesses have significant, positive correlations with observed employment, ranging 0.38 and 0.63 at the posterior median. Together both risk components account for between 38% and 48% of the observed variation at the posterior median. The table shows the historical variance decomposition. The first number in each cell is the posterior median. In parentheses, the 10th and 90th percentile are shown. The first row for each variable displays the fraction of the explained variance, while the second row lists the correlation of the counterfactual time series with the observed time series. The counterfactual assumes that only a subset of shocks were present. "All" shocks include also the aggregate real wage shock, the mean productivity shock, and initial conditions. Note that entry and average size and do not correspond exactly to the model implied definition, so their decomposition is not exact.

and 0.53 between the model and the data. The fraction of the variance explained by risk is 29 and 32%. At the firm-level, risk explains entry well, but often misses changes in average size. This low correlation may be due to a mismatch between the model and the data. In the model, exit occurs prior to production, unlike in the data. The model explains more than two thirds of observed entry. [38]

Note that risk shocks for young businesses have sizable effects on employment growth at older businesses, but they are largely uncorrelated with observed movements. This is unsurprising, given that in the data the correlation between employment of young at and older businesses is in the order of 0.2. The effect is sizable, since risk movements over the

[38] The variance decomposition varies little by sample period (see Tables C.3 and C.4 in the appendix).

full sample had a relative standard deviation of about one-half of the observed standard deviation. The correlation is negative and small, because the model mechanically implies a negative contemporaneous correlation due to general equilibrium price effects. Positive effects on older firms only manifest themselves over time as young firms age.

4.2.5 Relative risk or overall risk?

Does the distinction between upside and downside risk matter? To answer this question in more detail, I investigate a counterfactual in which the upside and downside risks are forced to move in unison. A counterfactual common overall risk factor can explain up to 50% to 70% of employment movements at young businesses, but it misses important movements in employment and exit rates.

To compute this counterfactual world with a constant risk composition, I use the historical shock decomposition again. Specifically, I adjust historical shocks so that their contribution to relative risk cancels, and only overall risk changes over time. The comparison between this counterfactual with the actual model estimates speaks to whether multiple risk components are needed to explain the data.[39]

Table 5: Fraction of variance of employment and exit accounted for by overall risk

	Employment age 1		Exit rate age 1	
Counterfactual	estab.	firms	estab.	firms
Baseline	0.36	0.45	1.00	1.00
Overall only, adjust both	-0.23	-0.09	0.79	0.55
Overall only, adjust downside	-0.33	-0.23	0.79	0.68
Overall only, adjust upside	0.26	0.23	0.32	-0.08

"Overall only" denotes a counterfactual that forces the ratio $\frac{\sqrt{\ell_t}}{\sqrt{r_t}}$, the relative risk, to be constant, while allowing overall risk $\sqrt{r_t^{-1}\ell_t^{-1}}$ to adjust. These counterfactual simulations that hold the relative risk constant explain employment and exit rates worse than the baseline estimate which allows both upside and downside risk to vary independently from each other. All results computed at the posterior mean.

Table 5 shows the implications of this experiment for employment and exit rates for businesses up to age one. At the estimated posterior mean, risk accounts for 36% of employment movements, and, by construction, virtually all the movements in the exit rate except initial conditions. Tying upside and downside risks together to keep overall risk constant decreases the fit significantly, across the specific counterfactuals.[40] In the

[39] Defining overall risk as the (geometric) average of the tail coefficients $1/\sqrt{\ell_t r_t}$ and relative risk as $\sqrt{\ell_t/r_t}$ allows to choose shocks and initial conditions to keep relative risk constant at zero.

[40] Because infinitely many linear combinations of shocks leave relative risk unchanged, I choose three particular combinations for their impact on overall risk. In each case shocks are rescaled to yield the same standard deviation of overall risk shocks. The first combination adjusts both shocks to r_t and ℓ_t, so that they offset each other in terms of relative risk. The second linear combination only adjusts downside risk ℓ_t, while the third only adjusts upside risk r_t. Initial conditions are chosen analogously.

first two scenarios, I adjust both risk components symmetrically or just upside risk. The resulting counterfactual misses employment dynamics completely, but it still explains most of the exit rates. The third scenario, which adjusts only downside risk, does better for employment, but it misses the exit rate completely.

4.2.6 Robustness

Changing the estimator for upside risk leaves the main results qualitatively unchanged. Figure C.13 in the Appendix compares the baseline specification to seven alternative specifications: (1) all variables are detrended prior to estimation, (2) both upside risk and employment are measured at age zero, (3) upside risk estimated at age one, (4) age zero risk is estimated using a GLS-type estimator, (5) risk is estimated using 20 employees as a cutoff compared to a baseline of 50 employees. Specifications (6) and (7) use job creation rather than employment. Appendix C.8 discusses the details. The median fraction of variance explained by risk is roughly constant around 40% across all specifications for establishment-level data. For firm-level data, the explained variance ranges from 29% to 56%.

4.3 Reduced form analysis

To assess whether my results are driven by particular industries or particular structural modeling assumptions, I estimate a statistical VAR model to assess the main model mechanisms. To identify shocks, I stop short of imposing all restrictions of the structural model – which would amount to estimating a VAR-approximation to the DSGE model – but instead rely on timing assumptions for shock-identification. I extract risk shocks from the VAR forecast errors by ordering upside risk first and the exit rate as a proxy for downside risk second. This implies that contemporaneously, only an upside risk shock affects upside risk and only the risk-shock and the exit shock affect the exit rate – a watered down version of the model implication of exogeneity of these two objects.[41] For comparison, I present results for the aggregate VAR first. In the interest of brevity, firm-level results are relegated to the Appendix.

Throughout, I estimate a VAR(2) with a flat Normal-Wishart prior, assuming disturbances are normally distributed, on the following variables: Employment at young businesses, exit rate of young businesses, entry (all in logs) as well as growth at old

The persistence of the shock processes is chosen as the average of the persistence of each component in the first case. Otherwise, the persistence of the unchanged risk component is chosen.

[41] Re-estimating the VAR with upside risk treated as exogenous to the other variables, which is a robust implication of my assumptions, does not change the IRFs significantly.

businesses at the same level of aggregation. At the industry level, I also add employment growth in the overall economy as a regressor. I also include a linear trend and constants.

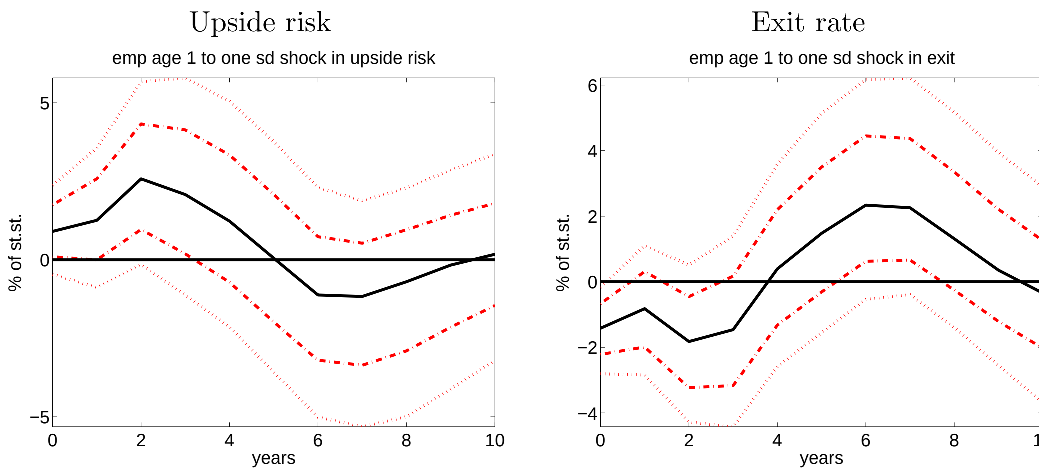


Figure 11: Upside risk, downside risk proxy, and employment at new establishments: VAR-based aggregate IRF

Shown are the median, 68, and 90% posterior probabilities of the IRF based on a Bayesian VAR. Shocks are identified using a Choleski-decomposition ordering upside risk and the exit rate first and second.

Figure 11 shows the responses of employment at young establishments to an upside risk shock and an exit rate shock. The 68%-posterior probability set is positive on impact for the upside risk shock and negative for the exit shock: Employment rises and falls in response to risk shocks as in the structural model (cf. Figure 9). Table C.1 in the Appendix shows that the qualitative results on impact are robust across different variable definitions. It also shows that average size rises with upside risk shocks and falls with exit shocks. There is some evidence that entry rises with upside risk shocks and robust evidence for entry falling in response to exit shocks.

I now turn to VARs estimated at the sectoral level to gauge which sectors are be driving my results. Table 6 summarizes the posterior median and 68% credible set of the impact response of employment of young establishments, entry, and average size. Table C.2 in the Appendix also provides the results at the firm-level. Employment and average size rises significantly in response to an upside risk shock in sectors 40 (utilities and telecommunication), 50 (wholesale), 60 (finance, insurance, real estate), and 70 (other services, including software). These sectors contribute, on average, about 55% of employment of young establishments.[42] The median response of entry to upside risk, however, is positive only in sectors 60 and 70, and insignificant. Exit rate shocks lowers

[42] At the firm-level, the results are also significantly positive in sector 70, as well as 20 (manufacturing) and 10 (mining), with most other results insignificant.

Table 6: Impact response to upside risk and exit rate shocks at new establishments: VAR-based sectoral results

SIC code	7	10	15-19	20-39	40-49	50-51	52-59	60-69	70-89
Employment	2.5	-1.3	-2.7	0.2	2.6	1.1	-0.4	4.1	2.4
to upside	(1.0,4.0)	(-4.6,2.0)	(-4.4,-1.0)	(-1.3,1.7)	(0.6,4.7)	(0.1,2.2)	(-1.3,0.5)	(2.6,5.8)	(1.1,3.7)
Entry	-2.2	-2.1	-1.2	-1.1	-0.7	-1.4	-1.1	0.4	0.5
to upside	(-2.9,-1.6)	(-3.3,-0.9)	(-2.1,-0.4)	(-1.7,-0.4)	(-1.6,0.2)	(-2.3,-0.5)	(-1.7,-0.6)	(-0.7,1.5)	(-0.3,1.3)
Avg size	4.7	0.8	-1.4	1.2	3.3	2.5	0.8	3.8	1.9
to upside	(3.4,6.2)	(-1.9,3.5)	(-2.6,-0.3)	(-0.0,2.5)	(1.4,5.4)	(1.6,3.4)	(0.0,1.5)	(1.7,5.9)	(0.7,3.1)
Employment	-3.2	-13.3	-6.1	-0.7	-1.7	-2.6	-1.0	-0.7	-0.9
to exit	(-4.6,-1.9)	(-16.1,-10.9)	(-7.6,-4.8)	(-2.2,0.7)	(-3.6,0.3)	(-3.6,-1.7)	(-1.9,-0.2)	(-2.1,0.8)	(-2.1,0.3)
Entry	-0.7	-3.8	-1.0	-1.0	-0.7	-0.7	-0.2	1.5	-2.0
to exit	(-1.3,-0.2)	(-4.9,-2.9)	(-1.8,-0.2)	(-1.6,-0.4)	(-1.5,0.2)	(-1.6,0.1)	(-0.7,0.3)	(0.5,2.5)	(-2.8,-1.4)
Avg size	-2.5	-9.4	-5.1	0.3	-1.0	-1.9	-0.8	-2.1	1.2
to exit	(-3.7,-1.4)	(-11.9,-7.3)	(-6.1,-4.3)	(-0.9,1.5)	(-2.9,0.9)	(-2.6,-1.2)	(-1.5,-0.1)	(-4.1,-0.2)	(0.1,2.3)
Emp. share	0.9	0.7	5.5	9.6	5.8	5.8	27.7	8.1	36.0

Shown are the posterior median and the 68% credible set in parentheses for a Bayesian VAR estimated sector by sector. Shocks are identified using a Choleski-decomposition ordering upside risk and the exit rate first and second. New business employment, exit, and risk are defined for businesses up to age one, as in the baseline aggregate specification. SIC classifications underlying the sectors: 7 (agriculture), 10 (mining), 15-19 (construction), 20-39 (manufacturing), 40-49 (transport and utilities), 50-51 (wholesale), 52-59 (retail), 60-69 (finance, insurance, real estate), 70-89 (business and other services).

entry significantly in most sectors, while the credible sets for employment and average size often include zero.[43]

Since the sectoral results for upside risk are strongest in sector 70, which includes software, and, at the establishment-level, for sector 40 at the establishment-level, which includes telecommunications, these results may suggest that the upside risk shock picked up newly available, but risky IT technology in the 1990s.

5 Conclusion

Employment at new businesses is important in the US economy. This paper provides a tractable framework for analyzing entrepreneurial entry and hiring in a dynamic general equilibrium framework. It provides conditions under which ex ante entrepreneurial risk can be identified from publicly available information on the size distribution and exit rates of new businesses.

Applied to US data from 1979 to 2011, the structural model implies that about 40% of time variation in employment at young businesses is explained by changes in the two components of entrepreneurial risk. This effect is sizable, but not unreasonably large: The remainder is explained by well-documented first moment shocks to productivity and real wages.

[43] The results for the exit rate at the firm-level are stronger and significant for almost all sectors.

There are two types of risk in my model: upside and downside risk. The novel approach here is to measure risk using data on the size distribution and exit rates for new businesses. The idea of the size distribution of the largest firms reflecting upside risk is more general than the specific model in this paper. Besides technical assumptions, it requires that among the most successful businesses, differences in size are mostly due to lucky draws from a risky productivity distribution. When this assumption is met, my approach to measurement allows me to recover ex ante risk from observed data.

Beyond its implications for job creation, my model can address other questions such as college and occupational choice by young agents when the payoff is risky. In addition, my model links movements in the average size of entrants to the risk premium investors demand. An open question is to what extent the estimated extent of crowding out through general equilibrium effects persists when the labor market is modeled in more detail. Exploring these implications is left for future research.

References

Adjemian, Stephane, Houtan Bastani, Michel Juillard, Ferhat Mihoubi, George Perendia, Marco Ratto, and Sebastien Villemot, "Dynare: Reference Manual, Version 4," Dynare Working Papers 1, CEPREMAP 2011.

Adrian, Tobias and Joshua Rosenberg, "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *Journal of Finance*, December 2008, *63* (6), 2997–3030.

Arkolakis, Costas, "A Unified Theory of Firm Selection and Growth," NBER Working Papers 17553, National Bureau of Economic Research October 2011.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz, "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy*, 2012, *120* (2), 304–345.

Blanchard, Olivier J, "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 1985, *93* (2), 223–47.

_ **and Jordi Galí**, "Real Wage Rigidities and the New Keynesian Model," *Journal of Money, Credit and Banking*, 2007, *39*, 35–65.

Bloom, Nicholas, "The Impact of Uncertainty Shocks," *Econometrica*, 2009, *77* (3), 623–685.

_ **, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry**, "Really Uncertain Business Cycles," 2011.

Bradford, William and Tatyana Sokolyk, "Ownership Structure and Performance of Closely Held Young Firms," 2012. mimeo, University of Washington.

Brooks, Stephen P and Andrew Gelman, "General methods for monitoring convergence of iterative simulations," *Journal of Computational and Graphical Statistics*, 1998, pp. 434–455.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *The Journal of Finance*, 2001, *56* (1), 1–43.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, February 2005, *113* (1), 1–45.

Clementi, Gian Luca and Berardino Palazzo, "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," Working Paper 19217, National Bureau of Economic Research July 2013.

_ **, Aubhik Khan, Berardino Palazzo, and Julia K. Thomas**, "Selection, Reallocation, and the Shape of Aggregate Fluctuations: A General Equilibrium Analysis," February 2013.

Cooley, Thomas F. and Edward C. Prescott, "Economic Growth and Business Cycles," in Thomas F. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton University Press, 1995, chapter 1, pp. 1–38.

Cooley, Thomas, Ramon Marimon, and Vincenzo Quadrini, "Aggregate Consequences of Limited Contract Enforceability," *Journal of Political Economy*, August 2004, *112* (4), 817–847.

DeBacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos, "The Properties of Income Risk in Privately Held Businesses," 2012. mimeo, Board of Governors of the Federal Reserve System.

Epstein, Larry G and Stanley E Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 1989, *57* (4), 937–69.

Fortin, Nicole M., "Higher-Education Policies and the College Wage Premium: Cross-State Evidence from the 1990s," *The American Economic Review*, 2006, *96* (4), pp. 959–987.

Garleanu, Nicolae and Stavros Panageas, "Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing," mimeo, Haas School of Business 2010.

Gomme, Paul and Peter Rupert, "Measuring labor's share of income," Policy Discussion Papers Nov, Federal Reserve Bank of Cleveland 2004.

Gordon, Robert J., "Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds," NBER Working Papers 18315, National Bureau of Economic Research August 2012.

Hall, Peter, "On Some Simple Estimates of an Exponent of Regular Variation," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1982, *44* (1), pp. 37–42.

Hall, Robert E., "The Stock Market and Capital Accumulation," *American Economic Review*, December 2001, *91* (5), 1185–1202.

_ , "Employment Fluctuations with Equilibrium Wage Stickiness," *American Economic Review*, March 2005, *95* (1), 50–65.

_ **and Susan E. Woodward**, "The Burden of the Nondiversifiable Risk of Entrepreneurship," *American Economic Review*, 2010, *100* (3), 1163–94.

Haltiwanger, John C., Ron S. Jarmin, and Javier Miranda, "Who Creates Jobs? Small vs. Large vs. Young," NBER Working Papers 16300, National Bureau of Economic Research, Inc August 2010.

Hansen, Lars P. and Thomas J. Sargent, "Recursive robust estimation and control without commitment," *Journal of Economic Theory*, September 2007, *136* (1), 1–27.

Heaton, John and Deborah Lucas, "Capital Structure: Hurdle Rates. and Portfolio Choice Interactions in an Entrepreneurial Firm," 2004. mimeo, Northwestern University.

Hill, Bruce M., "A Simple General Approach to Inference About the Tail of a Distribution," *The Annals of Statistics*, 1975, *3* (5), pp. 1163–1174.

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), pp. 1127–1150.

Hopenhayn, Hugo and Richard Rogerson, "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, October 1993, *101* (5), 915–38.

Hurst, Erik and Benjamin Wild Pugsley, "What Do Small Businesses Do?," *Brookings Papers on Economic Activity*, 2011, *103*, 73–142.

Jovanovic, Boyan and Peter L. Rousseau, "Stock Markets in the New Economy," in Chong-En Bai and Chi-Wa Yuen, eds., *Technology and the new economy*, MIT Press, 2002.

Judd, Kenneth L., "The law of large numbers with a continuum of IID random variables," *Journal of Economic Theory*, February 1985, *35* (1), 19–25.

Kelly, Bryan, "Dynamic Power Laws," 2012. mimeo, Booth School of Business.

Lee, Junghoon, "Does an Aggregate Increase in Idiosyncratic Volatility Cause a Recession?," 2011. mimeo, University of Chicago.

Lee, Yoonsoo and Toshihiko Mukoyama, "Entry, exit and plant-level dynamics over the business cycle," Working Paper 0718, Federal Reserve Bank of Cleveland 2008.

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

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

McGrattan, Ellen R. and Edward C. Prescott, "Unmeasured Investment and the Puzzling US Boom in the 1990s," *American Economic Journal: Macroeconomics*, October 2010, *2* (4), 88–123.

Moscarini, Giuseppe and Fabien Postel-Vinay, "The Timing of Labor Market Expansions: New Facts and a New Hypothesis," in "NBER Macroeconomics Annual 2008," Vol. 23, National Bureau of Economic Research, 2009, pp. 1–51.

Moskowitz, Tobias J. and Annette Vissing-Jørgensen, "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?," *American Economic Review*, September 2002, *92* (4), 745–778.

Oberfield, Ezra, "Business networks, production chains and productivity: A theory of input-output architecture," Technical Report 2011.

Pastor, Lubos and Pietro Veronesi, "Technological Revolutions and Stock Prices," *American Economic Review*, September 2009, *99* (4), 1451–83.

Prescott, Edward C., "Why do Americans work so much more than Europeans?," *Quarterly Review*, 2004, (Jul), 2–13.

Reed, William J. and Murray Jorgensen, "The Double Pareto-Lognormal Distribution – A New Parametric Model for Size Distributions," *Communications in Statistics – Theory and Methods*, 2004, *33* (8), 1733–1753.

Rossi-Hansberg, Esteban and Mark L. J. Wright, "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, December 2007, *97* (5), 1639–1666.

Schott, Immo, "Start-ups, House Prices, and the Jobless Recovery," 2013. EUI.

Shimer, Robert, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, March 2005, *95* (1), 25–49.

Uhlig, Harald, "A law of large numbers for large economies," *Economic Theory*, 1996, *8* (1), 41–50.

_ , "Explaining Asset Prices with External Habits and Wage Rigidities in a DSGE Model," *American Economic Review*, May 2007, *97* (2), 239–243.

Veracierto, Marcelo, "Firing Costs And Business Cycle Fluctuations," *International Economic Review*, 02 2008, *49* (1), 1–39.

Weil, Philippe, "Nonexpected Utility in Macroeconomics," *The Quarterly Journal of Economics*, February 1990, *105* (1), 29–42.

Welsh, A. H., "On the use of the empirical distribution and characteristic function to estimate parameters of regular variation," *Australian Journal of Statistics*, 1986, *28* (2), 173–181.

World Bank, "Doing Business 2013 – Smarter Regulations for Small and Medium-Sized Enterprises," Technical Report, World Bank 2013.

Appendix to "Entrepreneurial Tail Risk: Implications for Employment Dynamics"

A Proofs

A.1 Proof of Proposition 1: General solution properties

- (a) Sorting equilibrium

Part (a) of Proposition 1 asserts that the occupational choice problem has a sorting solution. This property stems from the fact that agents care only about wealth and are heterogeneous. Consider two entrepreneurs or self-employed with types O and O', $O > O'$. Since profits for entrepreneurs and for self-employed are proportional to productivity, type O earns strictly higher profits in every state than type O'. Given that the solution is interior and there is a continuum of agents, if agent O' enters, so must agent O. In equilibrium, the marginal agent is indifferent, but all agents of higher type strictly prefer entry.

Similarly, two entrepreneurs who have entered and ended up with productivity Z_0, Z_0' face a continuation payoff in each state that is strictly increasing in Z_0, Z_0'. If the entrepreneur with Z_0' exits, so should $Z_0 < Z_0'$ because continuation payoffs are smaller, state by state. Since agents entered to become entrepreneurs, there is some level of Z, say $\underline{Z}$, above which agents continue.

- (b) Aggregate production function:

Given that firms choose conditional on each state under full information, they solve a series of static profit maximization problems (3.17). For simplicity, I omit time and firm subscripts here.

Cost minimization implies that the capital-labor ratio is given by κ:

$$\kappa \equiv \frac{K}{N} = \frac{w}{r}\frac{\alpha}{1-\alpha}.$$

The firm therefore effectively solves $\max_N Z^{1-\phi} N^{\phi} \kappa^{\alpha\phi} - \frac{1}{1-\alpha} N w$, implying the following optimal labor demand:

$$N = Z\Big(\phi \frac{1-\alpha}{w}\Big)^{\frac{1}{1-\phi}} \kappa^{\frac{\alpha\phi}{1-\phi}} = Z \underbrace{\phi^{\frac{1}{1-\phi}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha\phi}{1-\phi}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha\phi}{1-\phi}}}_{\equiv \bar{N}(w,r)}.$$

Similarly:

$$K = Z \underbrace{\phi^{\frac{1}{1-\phi}} \left(\frac{1-\alpha}{w}\right)^{\frac{(1-\alpha)\phi}{1-\phi}} \left(\frac{\alpha}{r}\right)^{\frac{1-(1-\alpha)\phi}{1-\phi}}}_{\equiv \bar{K}(w,r)}.$$

Production is given by:

$$Y = Z \underbrace{\phi^{\frac{\phi}{1-\phi}} \Big(\frac{1-\alpha}{w}\Big)^{\frac{\phi}{1-\phi}(1-\alpha)} \Big(\frac{\alpha}{r}\Big)^{\frac{\phi}{1-\phi}\alpha}}_{\equiv \bar{Y}(w,r)}.$$

Profits are given by:

$$\Pi = (1-\phi)Y.$$

Conditional on survival, each firm's productivity grows at rate $g_{i,t}$. Since the death rate θ is *iid*, by a Law of Large Numbers, the total productivity of the entrepreneurial sector evolves as:

$$\begin{aligned} Z_{e,t} \equiv \int_{\mathcal{M}_t} Z_{e,i,t} di &= (1-\theta) \int_{\mathcal{M}_{t-1}} g_{i,t} Z_{e,i,t-1} di + \int_{\mathcal{M}_{0,t}} Z_{e,i,t} di \\ &= (1-\theta)\mathbb{E}[g_{i,t}] \int_{\mathcal{M}_{t-1}} Z_{e,i,t-1} di + Z_{e,0,t} = (1-\theta)\mathbb{E}[g_{i,t}] Z_{e,t-1} + Z_{e,0,t}, \end{aligned}$$

where $Z_{e,0,t}$ is the total productivity of entrants. Given linearity of policy functions in Z_{iat}, the firm-side of the model is equivalent to that of an aggregate entrepreneurial sector whose productivity follows the law of motion for $Z_{e,t}$.

- (c) Representative consumer:

Given that markets are complete and there is no arbitrage, individual risk is priced according to its physical probability, and aggregate risk is priced according to some stochastic discount factor $S_{t,t+\tau}$. The budget constraint conditional on survival can then be re-written as:

$$\begin{aligned} &\sum_{t=0}^{\infty} \int_{\zeta^t} \frac{Q_t(\zeta^t)}{Q_0(1-\theta)^t} \Big(B_{i,a-1,t-1}(\zeta^t) + A_{i,a-1,t-1}(\zeta^t)(P_t^s(t) + d_t^s(\zeta^t)) \\ &\qquad - (1-\theta)\big(P_t^s(\zeta^t) A_{iat}(\zeta^t) + B_{iat}(\zeta^t)\big)\Big) d\zeta^t \\ &\geq \sum_{t=0}^{\infty} \int_{\zeta^t} \frac{Q_t(\zeta^t)}{Q_0} \Big(C_{iat}(\zeta^t) - \pi_{iat}(\zeta^t)\Big) d\zeta^t \\ \Leftrightarrow & X_{i,0,t} \equiv \mathbb{E}_t \sum_{t=0}^{\infty} S_{0,t} \Big(\frac{B_{i,a-1,t-1} + A_{i,a-1,t-1}(P_t^s(t) + d_t^s) - (1-\theta)(P_t^s A_{iat} + B_{iat})}{(1-\theta)^t} + \pi_{iat} \Big) \\ &\geq \mathbb{E}_t \sum_{t=0}^{\infty} S_{0,t} C_{iat} \end{aligned}$$

where $X_{i,0,t}$ denotes total time 0 wealth of agent i and $B_{i,-1,t}(\cdot) = A_{i,-1,t} = 0$.

Given nonsatiation, the above budget constraint is binding in equilibrium. Let $\big(C^*_{i,a,t+a}(\zeta^{t+a}), B^*_{i,a,t+a}(\zeta^{t+a}\big)$ be the optimal policy of agent i born at time t with wealth $X_{i,0,t}$ given history ζ^{t+a}. Note that an agent with wealth $\tilde{X}_{i',0,t} = \lambda X_{i,0,t}$ can consume $\tilde{C}_{i',a,t+a}(\zeta_i^{t+a}) = \lambda C^*_{i,a,t+a}(\zeta_i^t)$ and similarly buy the proportionally scaled amount of securities and stocks state by state: This is budget-feasible. Given that preferences are homogeneous of degree one in consumption $C_{i,a,t+a}$, applying Euler's theorem shows that given maximization lifetime value must satisfy $V_{i,0,t} = \bar{v}_t X_{i,0,t}$ in equilibrium. More generally, repeating the above argument shows that in an interior equilibrium consumption and wealth scales state by state with initial wealth and so do continuation values. The following lemma summarizes these results:

Lemma 2. *In equilibrium, lifetime utility of agents born in the same cohort is proportional to their risk-adjusted lifetime wealth X_i:*

$$V_{iat} = \mathcal{R}[X_{iat}|\mathcal{F}_t^-]\bar{v}_t, \tag{A.1}$$

where $X_{i0t} = \mathbb{E}_t \sum_{s=0}^{\infty} S_{t,t+s}\pi_{i,s,t+s}$ is the present discounted value of lifetime nonfinancial income and unit wealth utility $\bar{v}_t > 0$.

Conditional on initial wealth X_{iat}, agents consume the same in every state of the world. Wealth of any two agents i, j at time t is proportional to initial relative wealth at time 0:

$$\lambda_{i,j} \equiv \frac{X_{i,0}}{X_{j,0}}, \quad V_{i,t} = \lambda V_{j,t},$$
$$C_{i,t} = \lambda C_{j,t}, A_{i,t} = \lambda A_{j,t}, B_{i,t} = \lambda B_{j,t}, X_{i,t} = \lambda X_{j,t}.$$

Consequently, there is a representative consumer in equilibrium, and the stochastic discount factor is given by:

$$S_{t+1} = \beta\Big(\frac{V_{t+1}}{\mathcal{R}(V_{t+1}|\mathcal{G}_t)}\Big)^{\psi-\rho}\Big(\frac{C_{o,t+1}}{C_{o,t}}\Big)^{-\psi}, \tag{A.2}$$

where $C_{o,t}$ denotes consumption of old agents (i.e. excluding the newborn).

A.2 Proof of Lemma 1

Part 1

The first part of Lemma 1, the characterization of the entry cutoff, follows from the characterization of lifetime utility in Lemma 2. Entrepreneurs enter as long as O satisfies:

$$\bar{v}_t PDV_t(W_t^f) \leq \mathcal{R}[\max\{(1-\eta)PDV_t(W_t^f)\bar{v}_t, Oe^{\mu_t}U_{i0t}PDV_t(G_e\bar{\Pi})\bar{v}_t\}|\mathcal{F}_{i0t}^-].$$

Note that $\mathcal{R}$ is homogeneous of degree one. Now, because there is no aggregate uncertainty at birth, $\bar{v}_t$ is $\mathcal{F}_{i0t}^-$ measurable and cancels in the equation, yielding for the marginal entrepreneur

$$1 = \mathcal{R}[\max\{(1-\eta), \bar{O}_{e,t}e^{\mu_t}\pi_t U_{i0t}\}|\mathcal{F}_{i0t}^-],$$

using the definition $\pi_t \equiv \frac{PDV_t(G_e\bar{\Pi}}{PDV_t(W^f)}$. Making the integral in $\mathcal{R}$ explicit, this equation becomes:

$$1 = \Big(\int_0^\infty \max\{(1-\eta), \bar{O}_{e,t}e^{\mu_t}\pi_t U_{i0t}\}^{1-\rho} dF_U(U_{i0t}|\ell_t, r_t)\Big)^{\frac{1}{1-\rho}}.$$

Thus, the risk-adjusted expectation is a function only of $\pi_t e^{\mu_t}$ and risk ℓ_t, r_t. Now guess and verify that $\bar{O}_{e,t} = \frac{g(\ell_t,r_t)}{\pi_t e^{\mu_t}}$ from equation (3.20) indeed solves the equation:

$$1 = \Big(\int_0^\infty \max\{(1-\eta), g(\ell_t, r_t)U_{i0t}\}^{1-\rho} dF_U(U_{i0t}|\ell_t, r_t)\Big)^{\frac{1}{1-\rho}}.$$

This equation implicitly defines $g(\ell_t, r_t)$.

Part 2

The second part of Lemma 1 uses the fact that the skill O is conditionally Pareto distributed if $\bar{O}_{e,t} > 1$. The unconditional distribution is given by

$$F_O(\tilde{O}) = \begin{cases} \frac{\nu}{\lambda+\nu}\tilde{O}^\lambda & \tilde{O} \le 1, \\ 1 - \frac{\lambda}{\lambda+\nu}\tilde{O}^{-\nu} & \tilde{O} > 1. \end{cases}$$

The associated conditional pdf is therefore given by

$$f_{O|O\ge\bar{O}}(\tilde{O}) = \frac{f_O(\tilde{O})}{1-F_O(\bar{O})} = \nu\frac{\frac{\lambda}{\lambda+\nu}\tilde{O}^{-\nu-1}}{\frac{\lambda}{\lambda+\nu}\bar{O}^{-\nu}} = \nu\left(\frac{\tilde{O}}{\bar{O}}\right)^{-\nu}\tilde{O}^{-1}, \tag{A.3}$$

with the conditional cdf given by $F_{O|O\ge\bar{O}}(\tilde{O}) = 1 - (\tilde{O}/\bar{O})^{-\nu}$.

Since entrepreneurs exit if $OUe^\mu\pi < (1-\eta)$, the exit rate conditional on skill O is given by $F_{U,t}\Big(\frac{1-\eta}{e^\mu \pi O}\Big)$. The overall exit rate is given by:

$$\begin{aligned}
EX_t &= \int_{\bar{O}_{e,t}}^\infty F_{U,t}\Big(\frac{1-\eta}{e^{\mu_t}\pi_t\tilde{O}}\Big)dF_O(\tilde{O}|O \ge \bar{O}_{e,t}) \\
&= \int_{\bar{O}_{e,t}}^\infty F_{U,t}\Big(\frac{1-\eta}{e^{\mu_t}\pi_t\tilde{O}}\Big)\Big(\frac{\tilde{O}}{\bar{O}_{e,t}}\Big)^{-\nu} d\ln(\tilde{O}) \\
&= \int_{\bar{O}_{e,t}}^\infty F_{U,t}\Big(\frac{1-\eta}{h_t(O)}\Big)\Big(\frac{h_t(O)}{h_t(\bar{O}_{e,t})}\Big)^{-\nu} d\ln(h_t(O)) \\
&= \int_{h_t(\bar{O}_{e,t})}^\infty F_{U,t}\Big(\frac{1-\eta}{x}\Big)\Big(\frac{x}{h_t(\bar{O}_{e,t})}\Big)^{-\nu} d\ln(x) \qquad \text{changing variables from } O \text{ to } x = h_t(O) \\
&= \int_{g(\ell_t,r_t)}^\infty F_{U,t}\Big(\frac{1-\eta}{x}\Big)\Big(\frac{x}{g(\ell_t,r_t)}\Big)^{-\nu} d\ln(x) \equiv \chi(\ell_t, r_t),
\end{aligned}$$

where $h_t(x) = xe^{\mu_t}\pi_t$ is strictly increasing and the last equality uses that $h_t(\bar{O}_{e,t}) = g(\ell_t, r_t)$ from the first part of the Lemma. The remaining integral is a function of ℓ_t, r_t only, because $F_{U,t}$ is a function of (ℓ_t, r_t) only:

$$F_{U,t}(x) = \begin{cases} \frac{r_t}{\ell_t+r_t}x^{\ell_t} & x \le 1, \\ 1 - \frac{\ell_t}{\ell_t+r_t}x^{-r_t} & x > 1. \end{cases}$$

A.3 Proof of Proposition 2: Connecting model and data

Part (b) of the Proposition follows from Lemma 1, equation (3.21), conditional on upside risk. It therefore remains to show that upside risk can be estimated.

Part (a) of the Proposition uses the following Theorem from Welsh (1986):

Lemma 3. *(Welsh, 1986, Theorem 2.1) Assume* $1 - F_Z(z) = \bar{F}_Z(z)$ *satisfies*

$$\bar{F}_Z(z) = \kappa_1 z^{-\omega}(1 + \kappa_2 z^{-(\nu-\omega)} + o(z^{-(\nu-\omega)})),$$

where $\kappa_1 > 0, \kappa_2 \neq 0, \omega > 0, \nu > \omega$. *Assume* $\{Z_i\}_{i=1}^n \overset{iid}{\sim} F_Z$, *with empirical distribution function* $F_n(z)$. *Let* $\infty > u > 1, \infty > m > 0$. *If* $z \propto m^{-1} n^{\frac{1}{2\nu-\omega}}$ *then as* $n \to \infty$*:*

$$n^{\frac{\nu-\omega}{2\nu-\omega}} \left(\frac{\ln(\bar{F}_n(uz)) - \ln(\bar{F}_n(z)}{-\ln(u)} - \omega \right)$$
$$\xrightarrow{d} \left(\frac{m^{-\frac{\omega}{2}}}{\ln(u)} \mathcal{N}\big(0, \frac{1}{\kappa_1}(1 - u^{-\omega})\big) + \frac{\kappa_2}{\kappa_1} \frac{m^{\nu-\omega}}{\ln(u)} (u^{\nu-\omega} - 1) \right).$$

To apply the theorem, I have to show that the density of productivity in my model gives rise to a Hall (1982) class distribution function that satisfies $\bar{F}_Z(z) = z^{-r}(\kappa_1 + \kappa_2 z^{-(\nu-r)} + o(z^{-(\nu-r)})), \kappa_1 > 0$.

Note that computing the density of the convolution of $\ln U$ and $\ln O$ gives the following untruncated density:

$$f_{\ln Z,t}(z) = \frac{r_t \ell_t}{r_t + \ell_t} \frac{\nu}{\nu - r_t} \begin{cases} e^{\ell_t \left(z - \ln \frac{g(r_t,\ell_t)}{\pi_t}\right)} \frac{\nu - r_t}{\nu + \ell_t} & e^z \leq \frac{g(r_t,\ell_t)}{\pi_t}, \\ e^{-r_t(z - \ln \frac{g(r_t,\ell_t)}{\pi_t})} \left(1 - \frac{\ell_t + r_t}{\nu + \ell_t} e^{-(\nu - r_t)(z - \ln \frac{g(r_t,\ell_t)}{\pi_t})}\right) & e^z > \frac{g(r_t,\ell_t)}{\pi_t}. \end{cases}$$

Direct integration of the density yields under Assumption 3:

$$\bar{F}_Z(z) = \begin{cases} \dfrac{1 - \frac{\nu}{\ell_t + \nu} \frac{r_t}{r_t + \ell_t} \left(\frac{z}{\bar{O}_{e,t}}\right)^{\ell_t}}{1 - \frac{\nu}{\ell_t + \nu} \frac{r_t}{r_t + \ell_t} \bar{U}_t(\bar{O}_{e,t})^{\ell_t}} & z \leq \bar{O}_{e,t} \\ \left(\frac{z}{\bar{O}_{e,t}}\right)^{-r_t} \frac{\ell_t}{\ell_t + r_t} \frac{\nu}{\nu - r} \dfrac{1 - \left(\frac{z}{\bar{O}_{e,t}}\right)^{-(\nu - r_t)} \frac{r_t}{\nu} \frac{\ell_t + r_t}{\ell_t + \nu}}{1 - \bar{U}_t(\bar{O}_{e,t})^{\ell_t} \frac{\nu}{\nu + \ell_t} \frac{r_t}{r_t + \ell_t}} & z > \bar{O}_{e,t} \end{cases}$$

Thus, for $Z > \frac{g(r_t,\ell_t)}{\pi_t}$ and $\nu > r$, $\bar{F}_Z(Z)$ belongs to the Hall class with $o(Z^{-(\nu-r)}) = 0$.

A.4 Proof of dynamic impact and comparative static effects in Table 1

A.4.1 Lemmas used in proof

I use the following two Lemmas to prove the results in Table 1.

Lemma 4. *Risk aversion lowers risk-adjusted expected value: For a nondegenerate random variable* $X > 0$*:* $\mathcal{R}_\rho[X] < \mathbb{E}[X]$ *iff* $\rho > 0$*. More generally, if* $\rho > \tilde{\rho}$*, then* $\mathcal{R}_\rho[X] < \mathcal{R}_{\tilde{\rho}}[X]$*.*

Let $g(x) = x^{1-\rho}$. If $\rho = 1$, $g(x) = \ln x$. Note that $\mathcal{R}[x] \equiv g^{-1}(\mathbb{E}[g(x)])$. If $\rho > 1$, $g'(x) < 0, g''(x) > 0$ and vice versa if $\rho \leq 1$.

$$g(\mathbb{E}[x]) < E[g(x)] \quad \Leftrightarrow \quad \mathbb{E}[x] \geq g^{-1}(E[g(x)]) \equiv \mathcal{R}[x],$$

where the first inequality is due to Jensen's inequality and the second inequality follows from the fact that g is decreasing and therefore its inverse g^{-1} is decreasing. Analogously for $\rho \leq 1$:

$$g(\mathbb{E}[x]) \geq E[g(x)] \quad \Leftrightarrow \quad \mathbb{E}[x] \leq g^{-1}(E[g(x)]) \equiv \mathcal{R}[x].$$

Assume $\rho < \tilde{\rho}$. Consider $\tilde{g}(x) = x^{1-\tilde{\rho}} = g(x)^{\frac{1-\tilde{\rho}}{1-\rho}}$. Define $h(x) = x^{\frac{1-\tilde{\rho}}{1-\rho}}$ and note that if $\rho < \tilde{\rho} < 1$ or $\tilde{\rho} > \rho > 1$, then $h'(x) = \frac{1-\tilde{\rho}}{1-\rho} x^{\frac{\rho-\tilde{\rho}}{1-\rho}} > 0, h''(x) = (1-\tilde{\rho})\frac{\rho-\tilde{\rho}}{(1-\rho)^2} x^{\frac{\rho-\tilde{\rho}}{1-\rho}-1} < 0$. If $\tilde{\rho} > 1 > \rho$ $h'(x) < 0, h''(x) > 0$. Applying the same reasoning as above with $g(x)$ taking the role of x shows that:

$$\tilde{\rho} > \rho \quad \Rightarrow \quad \mathcal{R}_{\rho}[x] \geq \mathcal{R}_{\tilde{\rho}}[x].$$

If x is not degenerate, the inequalities hold strictly.

Lemma 5. *First order stochastic dominance (FOSD): The distribution $F_{U,t}$ strictly dominates $F_{U,s}$ in the first order stochastic dominance sense if $\ell_t^{-1} < \ell_s^{-1}, \mu_t > \mu_s$ or $r_t^{-1} > r_s^{-1}$. Also, if $\ell_t r_t = \ell_s r_s$ while $\ell_t / r_t > \ell_s / r_s$, there is FOSD. That is, under these conditions*

$$F_{U,t}(u) < F_{U,s}(u), \qquad \forall u.$$

Proof: First, note that a since the logarithm is a strictly monotone transform, FOSD can be equivalently analyzed in levels or logs. Here it is convenient to work with logarithms. Second, since shifts in μ shift the location of the distribution in logs, FOSD is immediate. Write $\ln U_t - \mu_t = \mathcal{L}(\ell_t, r_t)$. Since the distribution function of the Laplace component is strictly increasing, it is immediate that $F_{\mathcal{L}(\ell_t,r_t)}(u - \mu_t) < F_{\mathcal{L}(\ell_t,r_t)}(u - \mu_s)$ iff $\mu_t < \mu_s$.

For simplicity. now consider $\mu_t = 0$. Note that:

$$F_{\mathcal{L}(\ell,r)}(x) = \begin{cases} \frac{r}{r+\ell} e^{\ell x} & x < 0, \\ 1 - \frac{\ell}{r+\ell} e^{-rx} & x \geq 0. \end{cases}$$

Assume $\tilde{\ell} < \ell, x < 0$. Then:

$$F_{\mathcal{L}(\ell,\ell)}(x) = \frac{r}{r+\ell} \exp(\ell x) < \frac{r}{r+\tilde{\ell}} \exp(\ell \omega x) < \frac{r}{r+\tilde{\ell}} \exp(\tilde{\ell} \omega x) = F_{\mathcal{L}(\tilde{\ell},r)}(x)$$

Assume $\tilde{\ell} < \ell, x \geq 0$. Then:

$$F_{\mathcal{L}(\ell,\ell)}(x) = 1 - \frac{r}{r+\ell} e^{-rx} < 1 - \frac{r}{r+\tilde{\ell}} e^{-rx} = F_{\mathcal{L}(\tilde{\ell},r)}(x)$$

Assume $\tilde{r} > r, x < 0$. Then:

$$F_{\mathcal{L}(\ell,r)}(x) = \frac{r}{r+\ell}\exp(\ell x) < \frac{\tilde{r}}{\tilde{r}+\ell}\exp(\ell x) = F_{\mathcal{L}(\ell,\tilde{r})}(x)$$

Assume $\tilde{r} > r, x \geq 0$. Then:

$$F_{\mathcal{L}(\ell,r)}(x) = 1 - \frac{\ell}{r+\ell}e^{-rx} < 1 - \frac{\ell}{\tilde{r}+\ell}e^{-rx} < 1 - \frac{\ell}{\tilde{r}+\ell}e^{-\tilde{r}x} = F_{\mathcal{L}(\ell,\tilde{r})}(x)$$

To prove the last case, define overall risk $\omega_t = \ell_t r_t$ and relative upside risk as $\upsilon = \frac{\ell_t}{r_t}$. Now consider the strictly monotone change of variables from $U \sim \mathcal{L}(\ell_t, r_t) = \omega_t \mathcal{L}(\upsilon_t, \upsilon_t^{-1})$ to $\frac{U}{\omega_t} \sim \mathcal{L}(\upsilon_t, \upsilon_t^{-1})$ for fixed ω_t. Since then the shift in relative risk is both a higher ℓ and a lower r, FOSD follows immediately from the above results.

A.4.2 Proof

- Dynamic impact effects

The proof of the effects of higher upside risk, lower downside risk, higher relative upside risk, or a higher mean log productivity μ on entry and overall productivity are based on first order stochastic dominance (FOSD, Lemma 5 on page VI) of the distribution $F_{U,t}$. Let $F_U \prec_{FOSD} \tilde{F}_U$. Defining $X(\tilde{U}) = \max\{(1-\eta)\bar{w}, \bar{O}_t\tilde{U}\bar{\Pi}\}$, the cutoff criterion can be written as:

$$\begin{aligned}
\bar{w} &= \Big(\int_0^\infty X(\tilde{U})^{1-\rho} dF_U(\tilde{U})\Big)^{\frac{1}{1-\rho}} \\
&= \Big(\big((1-\eta)\bar{w}\big)^{1-\rho}\big(1 - F_U(\bar{U}(\bar{O}))\big) + \int_{\bar{U}(\bar{O}))}^\infty (1 - F_U(\tilde{U}))dX(\tilde{U})^{1-\rho}\Big)^{\frac{1}{1-\rho}} \\
&< \Big(\big((1-\eta)\bar{w}\big)^{1-\rho}\big(1 - \tilde{F}_U(\bar{U}(\bar{O}))\big) + \int_{\bar{U}(\bar{O}))}^\infty (1 - \tilde{F}_U(\tilde{U}))dX(\tilde{U})^{1-\rho}\Big)^{\frac{1}{1-\rho}},
\end{aligned}$$

where the second inequality follows from the fact that $x^{1-\rho}$ is a decreasing function for $\rho \geq 1$, $(1 - \tilde{F}_U(\tilde{U})) > (1 - \tilde{F}_U(\tilde{U}))$, and the outer function $g(y) = y^{\frac{1}{1-\rho}}$ is decreasing for $\rho \geq 1$, interpreting the limit of $\rho = 1$ as natural log and exponential. For $\rho < 1$, the argument holds with the signs flipped.

Now note that the risk-adjusted expectation is strictly increasing in $\bar{O}$, since U is unbounded from above. Hence, the cutoff must be lower under $\tilde{F}_U$ than under F_U. Therefore, entry is higher.

Total productivity can be written as:

$$\int_{\bar{O}}^\infty \int_{\tilde{O}\tilde{U} \geq (1-\eta)\frac{\bar{w}}{\bar{\Pi}}} (1 - F_U(\tilde{U}))d\tilde{U}dF_O(\tilde{O}).$$

It increases because the integrand and the weight function are non-negative, the area of integration increases when $\bar{O}$ falls and the weight assigned to higher values in the distribution $(1 - F_U(\tilde{U}))$ increases with FOSD.

The effect of changing risk on average size is the result of opposing forces, the effect of worse types entering, but the underlying distribution improving. In general, the effects can go in either direction. Here I provide results for the limiting case of $\eta = 1$. By continuity, the results carry through for a sufficiently high η. Numerically, the results hold for values of η sufficiently below $\eta = 1$.

Generally, average size is given by:

$$\begin{aligned} &\bar{N}\frac{\bar{O}e^{\mu}}{(1-\ell^{-1})(1-r^{-1})(1-\nu^{-1})}\frac{(1+\frac{\ell}{r})(1+\frac{\ell}{\nu})-(1-\frac{1}{r})(1-\frac{1}{\nu})\bar{U}(\bar{O})^{1+\ell}}{(1+\frac{\ell}{r})(1+\frac{\ell}{\nu})-\bar{U}(\bar{O})^{\ell}} \\ &=\frac{\bar{N}\bar{\omega}}{\bar{\Pi}}\frac{g(\ell,r)}{(1-\ell^{-1})(1-r^{-1})(1-\nu^{-1})}\frac{(1+\frac{\ell}{r})(1+\frac{\ell}{\nu})-(1-\frac{1}{r})(1-\frac{1}{\nu})((1-\eta)/g(\ell,r))^{1+\ell}}{(1+\frac{\ell}{r})(1+\frac{\ell}{\nu})-((1-\eta)/g(\ell,r))^{\ell}} \end{aligned} \tag{A.4}$$

with $\bar{U}(\bar{O}) < 1$ as the exit threshold. The second line uses (3.20) from Lemma 1.

From equation (A.4) it is immediate that the average size is independent of the location parameter μ_t.

With infinite exit costs ($\eta = 1$), $\bar{U}(\bar{O}) = 0$ and the second factor drops out of equation A.4. In that case, the average size is proportional to $\frac{\mathbb{E}[U]}{\mathcal{R}[U]}$, which equals

$$\frac{\mathbb{E}[U]}{\mathcal{R}[U]} = \frac{1}{(1-1/r)(1+1/\ell)}\left(\frac{1}{(1-(1-\rho)/r)(1+(1-\rho)/\ell)}\right)^{-\frac{1}{1-\rho}}.$$

Direct differentiation shows the limiting results given in table 1 for downside risk ℓ, upside risk r, and relative upside risk.

To analyze overall risk, define overall risk as ω and relative risk as υ. For the effect of overall risk on entry, consider $\bar{O} \propto \frac{1}{\mathcal{R}[U]}$. Differentiation shows that:

$$\frac{d\bar{O}}{d\omega} \propto 1 - \upsilon^2 + 2\omega\upsilon(\rho - 1).$$

Thus if $1 - \upsilon^2 + 2\omega\upsilon(\rho - 1) < 0$, the entry threshold falls and entry rises.

The effect of overall risk on average size is positive if:

$$1 + \upsilon^2 - \omega\upsilon(\upsilon^2 - 1)(\rho - 2) + 2\omega^2\upsilon^2(\rho - 1) > 0.$$

It can be shown that $\frac{d\bar{O}}{d\omega} < 0$ (i.e. entry rises) is sufficient to guarantee the above condition.

- Comparative static effects

The comparative static effect of a higher outside option through higher wages on entry is immediate: it increases the payoff to entrepreneurs only when she exits (i.e. in only some states of the world) while decreasing profits in the remaining states of the world. Hence it increases the relative value of the outside option and lowers entry.

Now consider the average size of entrants in equation (A.4). The first term is the ratio of the labor share to the profit share and independent of wages. The second term is also independent of wages. Hence, average employment size is constant. It follows from the constant average size and lower number of entrants that total employment falls. In contrast, average productivity increases.

The comparative static effect of higher risk aversion follows because risk-adjusted expectations of a nondegenerate random variable strictly decrease in risk aversion ρ (Lemma 4 on page V). Thus entry falls with risk aversion. Total productivity falls as the minimum efficient scale is unchanged, but fewer entrepreneurs enter. However, average productivity and thus size increases under Assumption 3. To see this, differentiate (A.4) with respect to $\bar{O}$. The effect is proportional to:

$$
\begin{aligned}
&(\ell + r)(\ell + \nu) - \bar{U}(\bar{O})^{\ell}(r\nu(1+\ell) - \ell\bar{U}(\bar{O})(r-1)(\nu-1)) \\
&= (\ell + r)(\ell + \nu) - \bar{U}(\bar{O})^{\ell} r\nu(1+\ell) + \ell\bar{U}(\bar{O})^{1+\ell}(r-1)(\nu-1) \\
&> (\ell + r)(\ell + \nu) - \bar{U}(\bar{O})^{\ell} r\nu(1+\ell) \\
&> (\ell + r)(\ell + \nu) - 1(r\nu(1+\ell) = \ell + r + \nu - r\nu.
\end{aligned}
$$

The first inequality uses that $\ell\bar{U}(\bar{O})^{1+\ell} > 0$, the second inequality uses that $\ell\bar{U}(\bar{O})^{\ell} < 1$. A sufficient condition for higher risk aversion to increase average size is therefore $\ell + r + \nu - r\nu > 0$. However, also $\eta \nearrow 1$ or $\rho \to \infty$ is sufficient as it implies that $\bar{U}(\bar{O}) \to 0$.

B Additional equilibrium conditions

This section completes the model equations, first by detailing additional derivations not used in the previous proofs and then by summarizing all unknowns and equations.

B.1 Derivation of the number of entrants

Under the assumption that the entry cutoff $\bar{O} > 1$, the number of entrants is given by:

$$
\begin{aligned}
Entry_t &= \theta\epsilon \frac{\nu\lambda}{\nu+\lambda} \int_{\ln \bar{O}_{e,t}}^{\infty} \exp(-\nu \ln O) d\ln O \\
&= -\theta\epsilon \frac{\lambda}{\nu+\lambda}(0 - \bar{O}_{e,t}^{-\nu}) \\
&= \theta\epsilon \frac{1}{\nu\lambda^{-1}+1} \bar{O}_{e,t}^{-\nu} \qquad \text{(B.1)}
\end{aligned}
$$

B.2 Derivation of the exit rate

The exit rate is the probability of a low productivity draw, conditional on having a skill level O_{i0t} above the entry threshold $\bar{O}_{e,t}$. Different cases arise, depending on the location of the exit cutoff:

- Highest exit cutoff below the mode for productivity risk: $\frac{(1-\eta)}{\bar{O}_{e,t} e^{\mu_t} \pi_t} < 1$.

In this case, the exit rate is given by:

$$
\begin{aligned}
EX_t &= \int_{\bar{O}_{e,t}}^{\infty} F_{U,t}\left(\frac{(1-\eta)}{O_{i0t}\pi_t e^{\mu_t}}\right) dF_O(\tilde{O}|O \geq \bar{O}_{e,t}) \\
&= \int_{\bar{O}_{e,t}}^{\infty} \frac{r_t}{\ell_t + r_t}\left(\frac{(1-\eta)}{\pi_t \tilde{O}}\right)_t^{\ell} \nu \left(\frac{\tilde{O}}{\bar{O}_{e,t}}\right)^{-\nu} \tilde{O}^{-1} d\tilde{O} \\
&= \frac{\nu}{\nu + \ell_t}\frac{r_t}{\ell_t + r_t}\left(\frac{(1-\eta)}{\pi_t \bar{O}_{e,t}}\right)^{\ell_t}.
\end{aligned}
$$

Using that the product of the marginal entrepreneur's skills $\bar{O}_{e,t}$ times the profit ratio depends only on risk from (3.20) yields (3.21):

$$
EX_t = \frac{\nu}{\nu + \ell_t}\frac{r_t}{\ell_t + r_t}\left(\frac{(1-\eta)}{g(r_t, \ell_t)}\right)^{\ell_t} \equiv \chi(\ell_t, r_t). \tag{B.2}
$$

- Highest exit cutoff above the mode for productivity risk: $\frac{(1-\eta)}{\bar{O}_{e,t}e^{\mu_t}\pi_t} > 1$.

$$
\begin{aligned}
EX_t &= \int_{\bar{O}_t}^{\infty} F_{U,t}\left(\frac{(1-\eta)}{O_{i0t}\pi_t e^{\mu_t}}\right) dF_O(\tilde{O}|O \geq \bar{O}_{e,t}) \\
&= \int_{\frac{(1-\eta)}{e^{\mu_t}\pi_t}}^{\infty} \frac{r_t}{\ell_t + r_t}\left(\frac{(1-\eta)}{\pi_t \tilde{O} e^{\mu_t}}\right)^{\ell_t} \nu \left(\frac{\tilde{O}}{\bar{O}_{e,t}}\right)^{-\nu} \tilde{O}^{-1} d\tilde{O} \\
&+ \int_{\bar{O}_{e,t}}^{\frac{(1-\eta)}{e^{\mu_t}\pi_t}} \Big(1 - \frac{\ell_t}{\ell_t + r_t}\left(\frac{(1-\eta)}{\pi_t \tilde{O} e^{\mu_t}}\right)^{-r_t}\Big)\nu \left(\frac{\tilde{O}}{\bar{O}_{e,t}}\right)^{-\nu} \tilde{O}^{-1} d\tilde{O} \\
&= \frac{r}{\ell + r}\frac{\nu}{\nu + \ell}\Big(\frac{1-\eta}{\pi e^{\mu}}\Big)^{\ell} \bar{O}^{\nu}\Big(\frac{1-\eta}{\pi e^{\mu}}\Big)^{-(\nu+\ell)} \\
&+ 1 - \Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-\nu} - \frac{\nu}{\nu - r}\frac{\ell}{\ell + r}\bar{O}^{\nu}\Big(\frac{1-\eta}{\pi e^{\mu}}\Big)^{-r}\Big(\bar{O}^{-(\nu-r)} - \Big(\frac{1-\eta}{\pi O e^{\mu}}\Big)^{-(\nu-r)}\Big) \\
&= \frac{r}{\ell + r}\frac{\nu}{\nu + \ell}\Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-\nu} \\
&+ 1 - \Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-\nu} - \frac{\nu}{\nu - r}\frac{\ell}{\ell + r}\Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-r} + \frac{\nu}{\nu - r}\frac{\ell}{\ell + r}\Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-\nu} \\
&= 1 + \Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-\nu}\frac{\ell}{\nu - r}\frac{r}{\nu + \ell} - \frac{\nu}{\nu - r}\frac{\ell}{\ell + r}\Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-r} \\
&= 1 - \frac{\ell}{\nu - r}\Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-r}\Big(\frac{\nu}{\ell + r} - \frac{r}{\nu + \ell}\Big(\frac{1-\eta}{\pi \bar{O} e^{\mu}}\Big)^{-(\nu-r)}\Big)
\end{aligned} \tag{B.3}
$$

B.3 Derivation of the total productivity density

The density of log skills conditional on entry and $\bar{O} > 1$ is given by:

$$
f_{\ln O|O>\bar{O}}(o) = \frac{f_{\ln O}(o)}{1 - F_{\ln O}(\ln \bar{O})} = \frac{\frac{\nu\lambda}{\nu+\lambda}\exp(-\nu \ln O)}{Entry_t/\epsilon}.
$$

$$= \frac{\frac{\nu\lambda}{\nu+\lambda}\exp(-\nu o)}{\frac{\lambda}{\nu+\lambda}\bar{O}^{-\nu}} = \nu \exp(-\nu(o - \ln \bar{O})).$$

Actual log productivity is given by $\ln Z = \mu + \ln O + \ln U$. Without endogenous exit, the density is given by:

$$\begin{aligned}
f_{\ln Z}(z) &= \int_{\ln \bar{O}}^{\infty} f_{\ln O|O>\bar{O}}(o) f_{\ln U}(z - \mu - o) do \\
&= \mathbf{1}_{z-\mu>\ln \bar{O}} \int_{\ln \bar{O}}^{z-\mu} \nu \exp(-\nu(o - \ln \bar{O})) \frac{\ell r}{\ell + r} \exp(-r(z - \mu - o)) do \\
&+ \int_{\max\{z-\mu,\ln \bar{O}\}}^{\infty} \nu \exp(-\nu(o - \ln \bar{O})) \frac{\ell r}{\ell + r} \exp(\ell(z - \mu - o)) do \\
&= \mathbf{1}_{z-\mu>\ln \bar{O}} \frac{\ell r}{\ell + r} \exp(-r(z - \mu)) \nu \bar{O}^{\nu} \frac{1}{-(\nu - r)} \int_{\ln \bar{O}}^{z-\mu} -(\nu - r) \exp(-(\nu - r)o) do \\
&+ \nu \bar{O}^{\nu} \frac{\ell r}{\ell + r} \exp(\ell(z - \mu)) \frac{1}{-(\ell + \nu)} \int_{\max\{z-\mu,\ln \bar{O}\}}^{\infty} -(\ell + \nu) \exp(-(\nu + \ell)o) do \\
&= \mathbf{1}_{z-\mu>\ln \bar{O}} \frac{\ell r}{\ell + r} \exp(-r(z - \mu)) \nu \bar{O}^{\nu} \frac{1}{-(\nu - r)} \Big(\exp(-(\nu - r)(z - \mu)) - \exp(-(\nu - r) \ln \bar{O}) \Big) \\
&+ \nu \bar{O}^{\nu} \frac{\ell r}{\ell + r} \exp(\ell(z - \mu)) \frac{1}{-(\ell + \nu)} \Big(0 - \exp(-(\nu + \ell) \max\{z - \mu, \ln \bar{O}\}) \Big) \\
&= \mathbf{1}_{z-\mu>\ln \bar{O}} \frac{\ell r}{\ell + r} \exp(-r(z - \mu)) \nu \bar{O}^{\nu} \frac{\exp(-(\nu - r) \ln \bar{O})}{(\nu - r)} \Big(1 - \exp((\nu - r)(\ln \bar{O} - (z - \mu))) \Big) \\
&+ \nu \frac{\ell r}{\ell + r} \exp(\ell(z - \mu - \ln \bar{O})) \frac{\exp(-(\nu + \ell) \max\{z - \mu - \ln \bar{O}, 0\})}{(\ell + \nu)} \\
&= \mathbf{1}_{z-\mu>\ln \bar{O}} \frac{\ell r}{\ell + r} \exp(-r(z - \mu - \ln \bar{O})) \frac{\nu}{\nu - r} \Big(1 - \exp(-(\nu - r)(z - \mu - \ln \bar{O})) \Big) \\
&+ \nu \frac{\ell r}{\ell + r} \exp(\ell(z - \mu - \ln \bar{O})) \frac{\exp(-(\nu + \ell) \max\{z - \mu - \ln \bar{O}, 0\})}{(\ell + \nu)}
\end{aligned}$$

If $z - \mu \leq \ln \bar{O}$:

$$f_{\ln Z}(z) = \frac{\ell r}{\ell + r} \exp(\ell(z - \mu - \ln \bar{O})) \frac{\nu}{\ell + \nu}$$

If $z - \mu > \ln \bar{O}$:

$$\begin{aligned}
f_{\ln Z}(z) &= \frac{\ell r}{\ell + r} \exp(-r(z - \mu - \ln \bar{O})) \frac{\nu}{\nu - r} \Big(1 - \exp(-(\nu - r)(z - \mu - \ln \bar{O})) \Big) \\
&+ \nu \frac{\ell r}{\ell + r} \exp(\ell(z - \mu - \ln \bar{O})) \frac{\exp(-(\nu + \ell)(z - \mu - \ln \bar{O}))}{(\ell + \nu)} \\
&= \frac{\ell r}{\ell + r} \exp(-r(z - \mu - \ln \bar{O})) \frac{\nu}{\nu - r} \Big(1 - \exp(-(\nu - r)(z - \mu - \ln \bar{O})) \Big) \\
&+ \frac{\ell r}{\ell + r} \exp(-\nu(z - \mu - \ln \bar{O})) \frac{\nu}{(\ell + \nu)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\ell r}{\ell+r}\exp(-r(z-\mu-\ln\bar{O}))\frac{\nu}{\nu-r}\Big(1-\exp(-(\nu-r)(z-\mu-\ln\bar{O}))\Big(1-\frac{\nu-r}{\ell+\nu}\Big)\Big)\\
&= \frac{\ell r}{\ell+r}\exp(-r(z-\mu-\ln\bar{O}))\frac{\nu}{\nu-r}\Big(1-\exp(-(\nu-r)(z-\mu-\ln\bar{O}))\frac{\ell+r}{\ell+\nu}\Big)
\end{aligned}
$$

Since $\bar{O}=\frac{g(\ell,r)}{\pi e^{\mu}}$ it follows that $z-\mu-\ln\bar{O}=z-\ln\frac{g(\ell,r)}{\pi}$, with μ canceling out. The untruncated density can therefore be written as:

$$
f_{\ln Z}(z)=\frac{\ell r}{\ell+r}\frac{\nu}{\nu-r}\begin{cases} e^{\ell\left(z-\ln\frac{g(\ell,r)}{\pi}\right)}\frac{\nu-r}{\ell+\nu} & e^{z}\le\frac{g(\ell,r)}{\pi},\\ e^{-r\left(z-\ln\frac{g(\ell,r)}{\pi}\right)}\Big(1-e^{-(\nu-r)\left(z-\ln\frac{g(\ell,r)}{\pi}\right)}\frac{\ell+r}{\ell+\nu}\Big) & e^{z}>\frac{g(\ell,r)}{\pi}.\end{cases} \tag{B.4}
$$

For any $z\le\ln\bar{O}+\mu=\ln\frac{g(\ell,r)}{\pi}$, the untruncated CDF is given by:

$$
\begin{aligned}
F_{\ln Z}(z) &= \int_{-\infty}^{z} f_{lnZ}(\tilde{z})d\tilde{z}=\frac{\nu}{\nu+\ell}\frac{r}{r+\ell}\left(\bar{O}e^{\mu}\right)^{-\ell}[e^{\ell\tilde{z}}]_{-\infty}^{z}\\
&= \frac{\nu}{\nu+\ell}\frac{r}{r+\ell}\left(\frac{e^{z}}{\bar{O}e^{\mu}}\right)^{\ell}=\frac{\nu}{\nu+\ell}\frac{r}{r+\ell}\left(e^{z}\frac{\pi}{g(\ell,r)}\right)^{\ell}
\end{aligned}
$$

For $z>\ln\bar{O}+\mu$, the untruncated CDF is given by:

$$
\begin{aligned}
F_{\ln Z}(z) &= \int_{-\infty}^{\ln\bar{O}+\mu} f_{lnZ}(\tilde{z})d\tilde{z}+\int_{\ln\bar{O}+\mu}^{z} f_{lnZ}(\tilde{z})d\tilde{z}=\frac{\nu}{\nu+\ell}\frac{r}{r+\ell}\left(\frac{e^{\ln\bar{O}+\mu}}{\bar{O}e^{\mu}}\right)^{\ell}+\int_{\ln\bar{O}+\mu}^{z} f_{lnZ}(\tilde{z})d\tilde{z}\\
&= \frac{\nu}{\nu+\ell}\frac{r}{r+\ell}+\int_{\ln\bar{O}+\mu}^{z}\frac{\ell}{\nu-r}\Big(\frac{\nu}{\ell+r}re^{-r\left(\tilde{z}-\ln\bar{O}-\mu\right)}-\nu e^{-\nu\left(\tilde{z}-\ln\bar{O}-\mu\right)}\frac{r}{\ell+\nu}\Big)d\tilde{z}\\
&= \frac{\nu}{\nu+\ell}\frac{r}{r+\ell}+\frac{\ell}{\nu-r}\Big(\frac{\nu}{\ell+r}\Big(1-e^{-r\left(z-\ln\bar{O}-\mu\right)}\Big)-\frac{r}{\ell+\nu}\Big(1-e^{-\nu\left(z-\ln\bar{O}-\mu\right)}\Big)\Big)\\
&= \frac{\nu}{\nu+\ell}\frac{r}{r+\ell}+\frac{\ell}{\ell+r}\frac{\ell+\nu+r}{\ell+\nu}-\frac{\ell}{\nu-r}e^{-r\left(z-\ln\bar{O}-\mu\right)}\Big(\frac{\nu}{\ell+r}-\frac{r}{\ell+\nu}e^{-(\nu-r)\left(z-\ln\bar{O}-\mu\right)}\Big)\\
&= \frac{\nu r+\ell^{2}+\ell\nu+\ell r}{\nu r+\ell r+\ell^{2}+\ell\nu}-\frac{\ell}{\nu-r}e^{-r\left(z-\ln\bar{O}-\mu\right)}\Big(\frac{\nu}{\ell+r}-\frac{r}{\ell+\nu}e^{-(\nu-r)\left(z-\ln\bar{O}-\mu\right)}\Big)\\
&= 1-\frac{\ell}{\nu-r}e^{-r\left(z-\ln\bar{O}-\mu\right)}\Big(\frac{\nu}{\ell+r}-\frac{r}{\ell+\nu}e^{-(\nu-r)\left(z-\ln\bar{O}-\mu\right)}\Big)
\end{aligned}
$$

The density of equilibrium productivity is simply the above density truncated below at $\underline{z}=\ln\frac{(1-\eta)}{\pi}$ and given by the pdf divided by the complement of the exit rate given in (B.2) and (B.3). Similarly for the equilibrium cdf:

$$
f_{\ln Z}^{eqm}(z)=\frac{f_{\ln Z}(z)}{1-F_{\ln Z}(\underline{z})} \tag{B.5a}
$$

$$
F_{\ln Z}^{eqm}(z)=\frac{F_{\ln Z}(z)-F_{\ln Z}(\underline{z})}{1-F_{\ln Z}(\underline{z})} \tag{B.5b}
$$

Under the assumption that $\bar{U}=\frac{(1-\eta)}{Oe^{\mu}\pi}<1$ and using the definition that $\bar{O}=\frac{g(\ell,r)}{e^{\mu}\pi}$, it follows that $1-\eta < g(\ell, r)$. The truncation point $\underline{z} = \ln\frac{(1-\eta)}{\pi} < \ln\frac{g(\ell,r)}{\pi} = \ln\bar{O} + \mu$ is therefore in the left tail of the untruncated productivity density.

The exit rate is therefore given by

$$\int_{-\infty}^{\underline{z}} f_{lnZ}(z)dz = \frac{\nu}{\nu+\ell}\frac{r}{r+\ell}\left(\frac{1-\eta}{g(\ell,r)}\right)^{\ell},$$

which corresponds to equation (B.2).

B.4 Derivation of the expected total productivity

Total productivity of an entering cohort is the product of average productivity times the mass of entrants.

Average productivity of an entering cohort is given by:

$$\bar{Z}_{0,e} = EX \times 0 + (1-EX) \times \int_{\underline{z}}^{\infty} e^{\tilde{z}} f_z^{eqm}(\tilde{z})d\tilde{z}$$

using (B.5a) and assuming $\bar{U} < 1$ below

$$\begin{aligned}
&= \frac{1-EX}{1-EX}\frac{\ell r}{\ell+r}\frac{\nu}{\ell+\nu}e^{-\ell(\ln\bar{O}+\mu)}\int_{\underline{z}}^{\ln\bar{O}+\mu} e^{\tilde{z}(1+\ell)}d\tilde{z} \\
&+ \frac{1-EX}{1-EX}\frac{\ell r}{\ell+r}\frac{\nu}{\nu-r}e^{r(\ln\bar{O}+\mu)}\int_{\ln\bar{O}+\mu}^{\infty} e^{\tilde{z}(1-r)}d\tilde{z} - \frac{1-EX}{1-EX}\frac{\ell r}{\ell+\nu}\frac{\nu}{\nu-r}e^{\nu(\ln\bar{O}+\mu)}\int_{\ln\bar{O}+\mu}^{\infty} e^{\tilde{z}(1-\nu)}d\tilde{z} \\
&= \frac{\ell r}{\ell+r}\frac{\nu}{\ell+\nu}e^{-\ell(\ln\bar{O}+\mu)}\frac{1}{1+\ell}[e^{\tilde{z}(1+\ell)}]_{\underline{z}}^{\ln\bar{O}+\mu} \\
&+ \frac{\ell r}{\ell+r}\frac{\nu}{\nu-r}e^{r(\ln\bar{O}+\mu)}\frac{1}{r-1}[-e^{\tilde{z}(1-r)}]_{\ln\bar{O}+\mu}^{\infty} - \frac{\ell r}{\ell+\nu}\frac{\nu}{\nu-r}e^{\nu(\ln\bar{O}+\mu)}\frac{1}{\nu-1}[-e^{\tilde{z}(1-\nu)}]_{\ln\bar{O}+\mu}^{\infty}
\end{aligned}$$

using $1 < r < \nu$ below

$$\begin{aligned}
&= \frac{\ell r}{\ell+r}\frac{\nu}{\ell+\nu}e^{-\ell(\ln\bar{O}+\mu)}\frac{1}{1+\ell}(e^{(1+\ell)(\ln\bar{O}+\mu)} - e^{\underline{z}(1+\ell)}) \\
&+ \frac{\ell r}{\ell+r}\frac{\nu}{\nu-r}e^{r(\ln\bar{O}+\mu)}\frac{e^{-(r-1)(\ln\bar{O}+\mu)}}{r-1} - \frac{\ell r}{\ell+\nu}\frac{\nu}{\nu-r}e^{\nu(\ln\bar{O}+\mu)}\frac{e^{-(\nu-1)(\ln\bar{O}+\mu)}}{\nu-1} \\
&= \frac{\ell r}{\ell+r}\frac{\nu}{\ell+\nu}e^{\ln\bar{O}+\mu}\frac{1}{1+\ell}(1-e^{(\underline{z}-(\ln\bar{O}+\mu))(1+\ell)}) + \frac{\ell r}{\ell+r}\frac{\nu}{\nu-r}\frac{e^{\ln\bar{O}+\mu}}{r-1} - \frac{\ell r}{\ell+\nu}\frac{\nu}{\nu-r}\frac{e^{\ln\bar{O}+\mu}}{\nu-1} \\
&= \frac{\ell r}{\ell+r}\frac{\nu}{\nu-r}e^{\ln\bar{O}+\mu}\Big(\frac{\nu-r}{\ell+\nu}\frac{1}{1+\ell}(1-e^{(\underline{z}-(\ln\bar{O}+\mu))(1+\ell)}) + \frac{(\nu-r)(\ell-1)+(\nu+r)(\nu-r)}{(r-1)(\nu-1)(\ell+\nu)}\Big) \\
&= \frac{\ell r}{\ell+r}\frac{\nu}{\ell+\nu}e^{\ln\bar{O}+\mu}\Big(\frac{1}{1+\ell} + \frac{(\ell-1)+(\nu+r)}{(r-1)(\nu-1)} - \frac{1}{1+\ell}e^{(\underline{z}-(\ln\bar{O}+\mu))(1+\ell)}\Big) \\
&= \frac{r}{\ell+r}\frac{\nu}{\ell+\nu}\frac{\ell}{(1+\ell)}e^{\ln\bar{O}+\mu}\Big(\frac{(\ell+r)(\ell+\nu)}{(r-1)(\nu-1)} - e^{(\underline{z}-(\ln\bar{O}+\mu))(1+\ell)}\Big) \\
&= \frac{1}{1+\ell^{-1}}\frac{1}{1-r^{-1}}\frac{1}{1-\nu^{-1}}e^{\ln\bar{O}+\mu}\Big(1 - \frac{1-r^{-1}}{1+\ell r^{-1}}\frac{1-\nu^{-1}}{1+\ell\nu^{-1}}e^{(\underline{z}-(\ln\bar{O}+\mu))(1+\ell)}\Big)
\end{aligned}$$

using the definition of $\bar{U}$ below

$$= \frac{1}{1+\ell^{-1}}\frac{1}{1-r^{-1}}\frac{1}{1-\nu^{-1}}\bar{O}e^{\mu}\Big(1-\frac{1-r^{-1}}{1+\ell r^{-1}}\frac{1-\nu^{-1}}{1+\ell\nu^{-1}}\bar{U}^{1+\ell}\Big) \tag{B.6}$$

Combining the expression for average productivity (B.6) with the expression for the mass of entrants in (B.1) gives the total productivity of the entering cohort:

$$\begin{aligned} Z_{0,e} &= \bar{Z}_{0,e} \times Entry \\ &= \frac{\theta\epsilon\bar{O}^{-\nu}}{1+\nu\lambda^{-1}}\frac{\bar{O}e^{\mu}}{1+\ell^{-1}}\frac{1}{1-r^{-1}}\frac{1}{1-\nu^{-1}}\Big(1-\frac{1-r^{-1}}{1+\ell r^{-1}}\frac{1-\nu^{-1}}{1+\ell\nu^{-1}}\bar{U}(\bar{O})^{1+\ell}\Big) \end{aligned} \tag{B.7}$$

For ease of notation, subscripts have been suppressed in the derivation. In the expressions above, the exogenous parameters ℓ_t, r_t, μ_t and the endogenous cutoffs $\bar{O}_{e,t}, \bar{U}_t$ are time-varying.

B.5 Derivation of the expression for the risk-adjusted payoff

To compute the criterion for the risk adjusted payoff, simplify the notation by dropping time subscripts and writing W for $PDV(W)$ and Π for $PDV(\Pi)$:

$$W \le \mathcal{R}[\max\{(1-\eta)W, Oe^{\mu}U\Pi\}] = W\mathcal{R}[\max\{(1-\eta), Oe^{\mu}U(\Pi/W)\}]$$

$$= W\Big(\int_0^{\infty}\max\{(1-\eta)w, Oe^{\mu}\tilde{U}(\Pi/W)\}^{1-\rho}dF_U(\tilde{U})\Big)^{\frac{1}{1-\rho}}$$

$$= W\Big((1-\eta)^{1-\rho}\int_0^{\bar{U}}dF_U(\tilde{U}) + \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\int_{\bar{U}}^{\infty}\tilde{U}^{1-\rho}dF_U(\tilde{U})\Big)^{\frac{1}{1-\rho}} \qquad \text{where } \bar{U} \equiv (1-\eta)\frac{W}{Oe^{\mu}\Pi}$$

assuming $\bar{U} < 1$ below

$$\begin{aligned} = W\Big((1-\eta)^{1-\rho}&\int_{-\infty}^{\ln\bar{U}}\frac{\ell r}{\ell+r}\exp(\ell\ln\tilde{U})d\ln\tilde{U} \\ &+ \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\frac{\ell r}{\ell+r}\Big(\int_{\ln\bar{U}}^{0}\exp((\ell+1-\rho)\ln\tilde{U})d\ln\tilde{U} \\ &\qquad + \int_0^{\infty}\exp((-r+1-\rho)\ln\tilde{U})d\ln\tilde{U}\Big)\Big)^{\frac{1}{1-\rho}} \end{aligned}$$

$$= W\Big((1-\eta)^{1-\rho}\frac{r}{\ell+r}\bar{U}^{\ell} + \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\frac{\ell r}{\ell+r}\Big(\frac{1}{\ell+1-\rho}(1-\bar{U}^{\ell+1-\rho}) + \frac{1}{r+\rho-1}\Big)\Big)^{\frac{1}{1-\rho}}$$

collecting terms

$$= W\Big((1-\eta)^{1-\rho}\frac{r}{\ell+r}\bar{U}^{\ell} + \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\frac{\ell r}{\ell+r}\Big(\frac{1}{\ell+1-\rho}(-\bar{U}^{\ell+1-\rho}) + \frac{1}{\ell+1-\rho}\frac{r+\ell}{r+\rho-1}\Big)\Big)^{\frac{1}{1-\rho}}$$

using the definition of $\bar{U}$

$$= W\Big((1-\eta)^{1-\rho}\frac{r}{\ell+r}\bar{U}^{\ell} - \frac{\ell r}{\ell+r}\Big(\frac{\bar{U}}{1-\eta}\Big)^{\rho-1}\frac{1}{\ell+1-\rho}\bar{U}^{\ell+1-\rho} + \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\frac{1}{\ell+1-\rho}\frac{r\ell}{r+\rho-1}\Big)^{\frac{1}{1-\rho}}$$

$$= W\Big((1-\eta)^{1-\rho}\bar{U}^{\ell}\frac{r}{\ell+r}\Big(1-\frac{\ell}{\ell+1-\rho}\Big) + \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\frac{1}{1+(1-\rho)\ell^{-1}}\frac{1}{1-(1-\rho)r^{-1}}\Big)^{\frac{1}{1-\rho}}$$

using the definition of $\bar{U}$ again

$$= W\Big((1-\eta)^{1-\rho}\Big(\frac{(1-\eta)W}{Oe^{\mu}\Pi}\Big)^{\ell}\frac{r}{\ell+r}\frac{(1-\rho)\ell^{-1}}{1+(1-\rho)\ell^{-1}} + \Big(Oe^{\mu}\frac{\Pi}{W}\Big)^{1-\rho}\frac{1}{1+(1-\rho)\ell^{-1}}\frac{1}{1-(1-\rho)r^{-1}}\Big)^{\frac{1}{1-\rho}}$$

B.6 Capital and investment equilibrium conditions

The Law of Motion for capital is given by:

$$K_t = (1-\delta_t)K_{t-1} + (1-\Psi(\frac{I_t}{K_{t-1}}))I_t$$

The objective of the representative capital goods producing firm is given by:

$$\mathbb{E}_t\sum_{s=0}^{\infty}\frac{SDF_{t,t+s}}{(1-\theta)^s}(d_{t+s}K_{t+s} - I_{t+s}))$$

With P_t^k as the current value multiplier, maximizing the objective yields the following FOC:

$$[I] \qquad P_t(1-\Psi_t(\cdot) - \Psi_t'(\cdot)\frac{I_t}{K_{t-1}}) = 1$$

$$[K] \qquad d_t + \mathbb{E}_t[\frac{SDF_{t,t+1}}{1-\theta}((1-\delta_{t+1}) + \Psi_{t+1}'(\cdot)\left(\frac{I_t}{K_{t-1}}\right)^2)P_{t+1}] = P_t$$

The lifetime value can be express recursively as $V_t(k_{t-1})$. Note that the value $V_t(k_{t-1})$ is homogeneous of degree one. To see this, define $\iota \equiv \frac{i_t}{k_{t-1}}$. Then

$$\begin{aligned}
V_t(k_{t-1}) &= \max_{i_t} d_t\left((1-\delta)k_{t-1} + I_t(1-\Psi_t(\frac{I_t}{k_{t-1}}))\right) - I_t \\
&+ \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}V_{t+1}\left((1-\delta)k_{t-1} + I_t(1-\Psi_t(\frac{I_t}{k_{t-1}}))\right)\Big] \\
&\overset{\iota_t\equiv I_t/k_{t-1}}{=} \max_{\iota_t}\left(d_t((1-\delta) + \iota_t(1-\Psi_t(\iota_t))) - \iota_t\right)k_{t-1} \\
&+ \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}V_{t+1}\left(k_{t-1}(1-\delta) + k_{t-1}\iota_t(1-\Psi_t(\iota_t))\right)\Big] \\
&\overset{V_\tau(k_{\tau-1})\equiv P_\tau k_{\tau-1}}{=} k_{t-1}\max_{\iota_t}\left(d_t((1-\delta) + \iota_t(1-\Psi_t(\iota_t))) - \iota_t\right) \\
&+ \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}P_{t+1}\times((1-\delta) + \iota_t(1-\Psi_t(\iota_t))\Big] = P_t \times k_{t-1}
\end{aligned}$$

In competitive equilibrium, therefore, the marginal value of firm's capital stock is given by P_t, which is also simply the total value of the firm divided by its capital stock.

Now, $K_{t-1}\times P_t = V_t$ is the PDV of total dividends net of investment. Each agent gets $\zeta_t = d_t((1-\delta)+\iota_t(1-\Psi_t(\iota_t))) - \iota_t$ units of net dividends every period. Additionally, they can buy and sell shares at rate P_t, where

$$P_t = \zeta_t^* + \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}P_{t+1}\times((1-\delta) + \iota_t^*(1-\Psi_t(\iota_t^*))\Big]$$

The corresponding (state by state) budget constraint is given by:

$$C_t(i) + P_t\big(K_t(i) - \frac{K_{t-1}(i)}{1-\theta}(1-\delta_t + \iota_t(1-\Psi_t^*))\big) + \leq \omega_t(i) + d_t K_t(i) - \iota_t K_{t-1}(i)$$

The firm's FOCs above imply the household's no arbitrage condition holds:

$$\begin{aligned}
P_t - d_t &= \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}\Big(P_{t+1}\Big(1-\delta_{t+1}+\iota_{t+1}(1-\Psi_{t+1}^*)\Big) - \iota_{t+1}\Big)\Big] \\
&= \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}\Big(P_{t+1}(1-\delta_{t+1}) + \iota_{t+1}\Big(\underbrace{P_{t+1}(1-\Psi_{t+1}^*) - 1}_{=P_{t+1}\Psi'_{t+1}\iota_{t+1})}\Big)\Big)\Big] \\
&= \mathbb{E}_t\Big[\frac{SDF_{t,t+1}}{1-\theta}\Big(P_{t+1}(1-\delta_{t+1}) + (\iota_{t+1})^2\Psi'_{t+1}P_{t+1}\Big)\Big],
\end{aligned}$$

which equals the firms' FOC for capital. Thus, the firms' two FOCs imply the household's no arbitrage condition.

B.7 Derivation of Return on Wealth

B.7.1 Aggregate resource constraint

In what follows I use $\bar{c}(o)$ do denote the average per capita consumption of the old, as opposed to $c(o)$ which is the absolute amount.

Summing the flow budget constraints gives the flow resource constraint:

$$(1-\theta)\left(\bar{c}_t(o) + P_t\big(\bar{k}_t(o) - \frac{\bar{k}_{t-1}(o)}{1-\theta}(1-\delta_t+\iota_t(1-\Psi_t^*))\big)\right) \leq (1-\theta)\left(\omega_t(i) + d_t\bar{k}_t(o) - \iota_t\frac{\bar{k}_{t-1}(o)}{1-\theta}\right)$$

$$\theta\left(\bar{c}_t(y) + P_t\big(\bar{k}_t(y) - \frac{0}{1-\theta}(1-\delta_t+\iota_t(1-\Psi_t^*))\big)\right) \leq \theta\left(\omega_t(y) + d_t\bar{k}_t(y) - \iota_t 0 + \text{net end}_t\right)$$

$$\Rightarrow C_t + P_t\Big(\begin{Bmatrix}(1-\theta)\bar{k}_t(o)\\ +\theta\bar{k}_t(y)\end{Bmatrix} - \bar{k}_{t-1}(o)(1-\delta_t+\iota_t(1-\Psi_t^*))\Big) =$$

$$\begin{Bmatrix}(1-\theta)\bar{\omega}_t(o)\\ +\theta\bar{\omega}_t(y)\end{Bmatrix} + d_t\begin{Bmatrix}(1-\theta)\bar{k}_t(o)\\ +\theta\bar{k}_t(y)\end{Bmatrix} - \iota_t\bar{k}_{t-1}(o) + \theta\text{net end}_t$$

$$\Leftrightarrow C_t + P_t\underbrace{\left(K_t - K_{t-1}(1-\delta_t+\iota_t(1-\Psi_t^*))\right)}_{=0} = \omega_t + d_tK_t - I_t + \theta\text{net end}_t$$

which coincides with the resource constraint

$$C_t = \omega_t + dK_t - I_t + \theta\text{net end}_t,$$

where ω_t is the total human capital income.

In steady state $S^* = 0$, $I = \delta K$:

$$C = \omega + (d-\delta)K + \theta\text{net end}_t.$$

Below restrictions on d are derived.

B.7.2 Preferences, SDF, and change of measure

Preferences:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \tilde{\beta}^t \mathbf{1}_{\text{Survival},0\to t} \ln c_t = \sum_{t=0}^{\infty} (\tilde{\beta}(1-\theta))^t \ln c_t = \sum_{t=0}^{\infty} \beta^t \ln c_t,$$

or (with recursive utility):

$$\ln V_t = (1-\beta)\ln C_t + \beta \ln \mathcal{R}_t[V_{t+1}], \quad \mathcal{R}_t[X_{t+1}] = \mathbb{E}_t[X_t^{1-\gamma}]^{\frac{1}{1-\gamma}},$$

so that the SDF is given by:

$$\ln v_t(X_t) = (1-\beta)\ln c_t + \beta \ln \mathcal{R}_t[\exp(\ln v_{t+1}(X_{t+1}))] \tag{B.8}$$

Guess that $v_t(X_t) = \bar{v}_t \times X_t$ and $c_t = (1-\beta)X_t$. Then $X_{t+1} = R^w_{t+1}\beta X_t$:

$$\ln \bar{v}_t + \ln X_t = (1-\beta)\ln(1-\beta) + (1-\beta)\ln X_t + \beta \ln(\beta X_t) + \beta \ln \mathcal{R}_t[e^{\ln \bar{v}_{t+1} + \ln R^w_{t+1}}].$$

Canceling:

$$\ln \bar{v}_t = (1-\beta)\ln(1-\beta) + \beta \ln(\beta) + \beta \ln \mathcal{R}_t[e^{\ln \bar{v}_{t+1} + \ln R^w_{t+1}}].$$

Marginal utility of consumption today:

$$MPC_t = (1-\beta)c_t^{-1}.$$

In level terms:

$$\frac{\partial v_t}{\partial c_t} = (1-\beta)\frac{v_t}{c_t}$$

Benefit of increasing consumption tomorrow $\mathcal{R}_t[x] = \mathbb{E}_t[x^{1-\gamma}]^{\frac{1}{1-\gamma}}$.

$$\beta \frac{v_t}{\mathcal{R}[\cdot]} \frac{1}{1-\gamma} \mathcal{R}_t[v_{t+1}]^{\gamma} \mathbb{E}_t[(1-\gamma)v_{t+1}^{-\gamma}(1-\beta)\frac{v_{t+1}}{c_{t+1}}]$$

Thus the SDF is given by:

$$SDF_{t,t+1} = \beta \left(\frac{v_{t+1}}{\mathcal{R}_t[v_{t+1}]} \right)^{1-\gamma} \frac{c_t}{c_{t+1}}$$

$$\mathcal{R}_t[v_{t+1}] = \left(\frac{v_t}{c_t^{1-\beta}} \right)^{\frac{1}{\beta}} = \left(\frac{\bar{v}_t X_t}{((1-\beta)X_t)^{1-\beta}} \right)^{\frac{1}{\beta}} = \left(\frac{\bar{v}_t}{(1-\beta)^{1-\beta}} \right)^{\frac{1}{\beta}} X_t$$

Thus:

$$SDF_{t,t+1} = \beta \left(\frac{\bar{v}_{t+1} X_{t+1}}{\left(\frac{\bar{v}_t}{(1-\beta)^{1-\beta}}\right)^{1/\beta} X_t} \right)^{1-\gamma} \frac{c_t}{c_{t+1}} = \beta \left(\frac{\bar{v}_{t+1} \beta R^w_{t+1}}{\left(\frac{\bar{v}_t}{(1-\beta)^{1-\beta}}\right)^{1/\beta}} \right)^{1-\gamma} \frac{c_t}{c_{t+1}}$$

And $\frac{X_{t+1}}{X_t \beta} = R^w_{t+1}$ or $\frac{X_{t+1}}{X_t} = \beta R^w_{t+1}$

$$SDF_{t,t+1} = \beta \frac{C_t}{C_{t+1}} \left(\frac{V_{t+1}}{\mathcal{R}_t[V_{t+1}]} \right)^{1-\gamma} = \beta \frac{C_t}{C_{t+1}} \left(\frac{V_{t+1}}{(V_t/C_t)^{1/\beta} C_t} \right)^{1-\gamma} = \beta \left(\frac{C_t}{C_{t+1}} \right)^{\gamma} \left(\frac{V_{t+1}/C_{t+1}}{(V_t/C_t)^{1/\beta}} \right)^{1-\gamma}$$

Note that the SDF can be written as $\beta \frac{C_t}{C_{t+1}}$ times the martingale $\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}^{\mathbb{P}}_t[V_{t+1}^{1-\gamma}]}$, which represent the change of probability measure from the physical probability measure $\mathbb{P}$ to the (partially) risk-adjusted probability measure $\mathbb{Q}$: $\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t[V_{t+1}^{1-\gamma}]}$.

B.7.3 Equilibrium consumption share and return to wealth

Household optimization is subject to the budget constraint for old agents (which holds state by state due to complete markets, but whose state dependence is suppressed for notational simplicity):

$$\bar{c}_t(o) + P_t\big(\bar{k}_t(o) - \frac{\bar{k}_{t-1}(o)}{1-\theta}(1 - \delta_t + \iota_t(1 - \Psi^*_t))\big) \leq \omega_t(i) + d_t \bar{k}_t(o) - \iota_t \frac{\bar{k}_{t-1}(o)}{1-\theta}$$

As discussed above, in equilibrium (when the FOC w.r.t. investment holds), this implies the following FOC:

$$P_t - d_t = \mathbb{E}_t \frac{SDF_{t,t+1}}{1-\theta} P_{t+1}(1 - \delta_{t+1} + \iota^2_{t+1}\Psi'_{t+1})$$

For future reference, defining $\tilde{\delta}_t \equiv \delta_t - (\iota_t)^2 \Psi'_t$[44]:

$$\begin{aligned} P_t - d_t &= \mathbb{E}^{\mathbb{P}}_t[SDF_{t,t+1} P_{t+1} \frac{1 - \tilde{\delta}_{t+1}}{1-\theta}] \\ &= \mathbb{E}^{\mathbb{Q}}_t[\beta \frac{C_t}{C_{t+1}} P_{t+1} \frac{1 - \tilde{\delta}_{t+1}}{1-\theta}] \\ \frac{1-\theta}{1-\tilde{\delta}_{t+1}} \frac{P_t - d_t}{P_{t+1}} &= \frac{\beta W_t}{W_{t+1}} = (R^w_{t+1})^{-1} \\ R^w_{t+1} &= \frac{1 - \tilde{\delta}_{t+1}}{1-\theta} \frac{P_{t+1}}{P_t - d_t} = \frac{\text{Return on capital}_{t+1}}{1-\theta} \end{aligned}$$

Using the FOC for investment, the BC can be re-written as:

$$\bar{c}_t(o) + P_t\big(\bar{k}_t - \frac{\bar{k}_{t-1}}{1-\theta}(1 - \delta_t + (\iota_t)^2 \Psi'_t\big) \leq \omega_t(i) + d_t \bar{k}_t(o)$$

[44] Since Ψ' is an increasing function, the effective depreciation rate falls in k if $\iota > \delta$ (if investing more tomorrow, it is good to have a higher capital stock today) and rises otherwise.

Write the explicit complete markets BC:

$$\sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \mathbb{E}_t^{\mathbb{Q}} \Big[\beta^t \frac{C_0}{C_t} \Big(c_t + P_t \big(\bar{k}_t(o) - \frac{\bar{k}_{t-1}(o)}{1-\theta} (1 - \tilde{\delta}_t) \big) - (\omega_t(i) + d_t k_t) \Big) \Big] \leq 0$$

Note that since individual consumption is proportional to aggregate consumption state by state and there is non-satiation:

$$\begin{aligned}
\frac{c_0}{1-\beta} &= \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \mathbb{E}_t^{\mathbb{Q}} \Big[\beta^t \frac{C_0}{C_t} \Big(P_t \big(\frac{\bar{k}_{t-1}}{1-\theta} (1 - \tilde{\delta}_t) - \bar{k}_t \big) + (\omega_t(i) + d_t k_t) \Big) \Big] \\
&= P_0 \frac{\bar{k}_{-1}}{1-\theta} (1 - \tilde{\delta}_0) + \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \beta^t \frac{C_0}{C_t} \mathbb{E}_t^{\mathbb{Q}} \Big[\omega_t(i) \Big] \\
&\quad + \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \Big[\beta^t \frac{C_0}{C_t} k_t \times \underbrace{\mathbb{E}_t^{\mathbb{Q}} \Big[\Big(\beta \frac{C_t}{C_{t+1}} P_{t+1} \frac{1 - \tilde{\delta}_{t+1}}{1-\theta} (1 - \tilde{\delta}_{t+1}) - (P_t - d_t) \Big) \Big]}_{=0 \text{ by no arbitrage / FOC}} \Big] \\
&= P_0 \frac{\bar{k}_{-1}}{1-\theta} (1 - \tilde{\delta}_0) + \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \beta^t \frac{C_0}{C_t} \mathbb{E}_t^{\mathbb{Q}} \Big[\omega_t(i) \Big] \qquad \text{(B.9)}
\end{aligned}$$

Thus the consumption share of lifetime wealth is $(1 - \beta)$ and lifetime wealth is given by the value of the capital holdings plus the PDV of human income.

$$\begin{aligned}
\text{Lifetime wealth}_t &= P_0 \frac{\bar{k}_{-1}}{1-\theta} (1 - \tilde{\delta}_0) + \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \beta^t \frac{C_0}{C_t} \mathbb{E}_t^{\mathbb{Q}} \Big[\omega_t(i) \Big] \\
&= P_0 \frac{\bar{k}_{-1}}{1-\theta} (1 - \tilde{\delta}_0) + \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \left(\prod_{u=1}^{s} (R^w_{t-1+u})^{-1} \right) \mathbb{E}_t^{\mathbb{Q}} \Big[\omega_t(i) \Big]
\end{aligned}$$

To compute the evolution of the return on wealth from aggregate variables note that the wealth in the economy is proportional to consumption: $W_t = (1 - \beta)^{-1} C_t$. The wealth of the old is given by the overall wealth in the economy, minus the wealth of the young. The wealth of the young is given by:

$$W_t^y = \theta \overline{\text{net end}}_t + \theta \sum_{t=0}^{\infty} \mathbb{E}_0^{\mathbb{Q}} \left(\prod_{u=1}^{s} (R^w_{t-1+u})^{-1} \right) \mathbb{E}_t^{\mathbb{Q}} \Big[\bar{\omega}_t(y) \Big] \equiv \theta \Big(\overline{\text{net end}}_t + \overline{PDV(\omega(y))} \Big)$$

Thus:

$$R^w_{t-1,t} = \frac{W_t^o}{(1 - (1 - \beta)) W_{t-1}^o} = \frac{W_t - W_t^y}{\beta (1 - \theta) W_{t-1}} \qquad \text{(B.10)}$$

In steady state

$$c_t(o) = (1-\beta) \frac{1-\delta}{1-\theta} P_t k_{t-1}(o) + (1-\beta) \frac{1}{1 - (\bar{R}^w)^{-1}} w(o) = (1-\beta) \frac{1-\delta}{1-\theta} P_t k_{t-1}(o) + (1-\beta) \frac{\bar{R}_w}{\bar{R}_w - 1} w(o)$$

The expression for young agents is the same, except for an added endowment term:

$$c_t(y) = (1-\beta)\Big(\frac{\bar{R}_w}{\bar{R}_w - 1} w(y) + \text{net end}_t\Big)$$

Summing over all agents must gives in steady state

$$\begin{aligned} C_t &= (1-\beta)(1-\delta)P_t\bar{k}_{t-1}(o)(1-\theta) + (1-\beta)PDV(\omega) + \theta\text{net end} \\ &= (1-\beta)(1-\delta)P_t\bar{K}_{t-1} + \frac{(1-\beta)\bar{R}_w}{\bar{R}_w - 1}\omega + (1-\beta)\theta\text{net end} \end{aligned}$$

Compute $\bar{d}$ to make the budget constraint and the resource constraint hold:

$$C = (d-\delta)K + \omega + \theta\text{net end} = (1-\beta)(1-\delta)K + \frac{(1-\beta)\bar{R}_w}{\bar{R}_w - 1}\omega + (1-\beta)\theta\text{net end}$$

Hence

$$\big((d-\delta) - (1-\beta)(1-\delta)\big)K + \theta\beta\text{net end} = \frac{(1-\beta)\bar{R}_w - (\bar{R}_w - 1)}{\bar{R}_w - 1}\omega = \frac{1-\beta\bar{R}_w}{\bar{R}_w - 1}\omega$$

From the EE $1 - d = (\bar{R}^w)^{-1}\frac{1-\delta}{1-\theta}$ and $d - \delta = (1-\delta)(1-(\bar{R}^w)^{-1}(1-\theta)^{-1})$. Then:

$$(1-\delta)(\beta - ((1-\theta)\bar{R}_w)^{-1})K + \theta\beta\text{net end} = \frac{1-\beta\bar{R}_w}{\bar{R}_w - 1}\omega$$

Without mortality risk, $\bar{R}^w = \beta^{-1} = \tilde{\beta}^{-1}$, and the equation holds trivially.

If $\bar{R}^w = ((1-\theta)\beta)^{-1} > \beta^{-1}$, then the LHS is weakly positive. However, the RHS is negative at this $\bar{R}^w$ because $\beta\bar{R}^w = (1-\theta)^{-1} > 1$. Thus, the return on wealth has to be lower.

A lower return on wealth lowers the LHS, turning it negative eventually, while leaving the RHS negative. Note that at $\bar{R}^w = \beta^{-1} > \tilde{\beta}^{-1}$ the RHS is zero, while the LHS is negative for θ(net end) small enough. By continuity, therefore:

$$\bar{R}^w \in (\beta^{-1}, \beta^{-1}(1-\theta)^{-1}) \text{ if net end} \searrow 0.$$

Since d is increasing in $\bar{R}^w$, the marginal product of capital is higher in an economy with mortality risk.

Rewrite the above:

$$0 \le \theta\beta\frac{\text{net end}}{K} = \frac{1-\beta\bar{R}_w}{\bar{R}_w - 1}\frac{\omega}{K} + (1-\delta)\left(\frac{1}{(1-\theta)\bar{R}_w} - \beta\right)$$

Lemma 6. *There is a unique solution $\bar{R}_w \in (1, \beta^{-1}(1-\theta)^{-1})$ for general net end ≥ 0 and $\bar{R}_w \in (\beta^{-1}, \beta^{-1}(1-\theta)^{-1})$ if net end $= 0$.*

Proof: Note that the RHS is strictly decreasing in R_w for $R_w > 1$ because $\frac{1-\frac{d}{dR_w}\beta\bar{R}_w}{R_w - 1} = \frac{-(1-\beta)}{(R_w-1)^2}$ and R_w^{-1} is trivially decreasing. The function approaches $+\infty$ as $R_w \searrow 1$ and turns negative as $R_w \nearrow \beta^{-1}(1-\theta)^{-1}$. Since the LHS is non-negative, there is therefore a unique solution $\bar{R}_w \in (1, \beta^{-1}(1-\theta)^{-1})$. If net end $= 0$, then the solution satisfies $\bar{R}_w \in (\beta^{-1}, \beta^{-1}(1-\theta)^{-1})$.

Note that if the net endowment is large enough and since the return on wealth is higher than $\beta^{-1} = (\tilde{\beta}(1-\theta))^{-1}$, the young do not fully consume their labor income but instead accumulate capital because their propensity to consume out of human income is given by:

$$\frac{1-\beta}{1-1/\bar{R}_x} < 1.$$

If the net endowment of young agents is very high, this lowers the overall return on wealth as the young run down their savings in a consumption spree.

B.8 Full model equations

Unknowns: (1) capital stock K_t, (2) rental rate of capital d_t^k, (3) price of capital P_t^k, (4) investment I_t, (5) cost wage W_t, (6) effective wage W_t^f, (7) entry cutoff for entrepreneurs $\bar{O}_{e,t}$, (8) exit rate EX_t, (9) productivity of entering entrepreneurs $Z_{0,t}$, (10) overall productivity of entrepreneurs $Z_{e,t}$, (11) unit labor demand $\bar{N}_t$, (12) unit capital demand $\bar{K}_t$, (13) stochastic discount factor S_t, (14) present discounted value of unit profits $PDV_t(\Pi)$, (15) present discounted value of unit wages $PDV_t(W^f)$, (16) present discounted value of wages and home production for nonentrepreneurs $PDV_t(W^+)$, (17) effective labor supply of potential entrepreneurs $N_{e,t}^s$, (18) consumption, (19) total output, (20) unit profits, (21) return on wealth, (22) unit value function.

$$(1)\ K_t = (1-\delta_t)K_{t-1} + I_t(1-\Psi(I_t/K_{t-1}))$$

$$(2)\ K_t = Z_{e,t}\bar{K}_t$$

$$(3)\ \bar{K}_t = \phi^{\frac{\phi}{1-\phi}}\Big(\frac{1-\alpha}{W_t}\Big)^{\frac{\phi}{1-\phi}(1-\alpha)}\Big(\frac{1-\alpha}{d_t^k}\Big)^{\frac{1-(1-\alpha)\phi}{1-\phi}}$$

$$(4)\ \bar{N}_t = \phi^{\frac{\phi}{1-\phi}}\Big(\frac{1-\alpha}{W_t}\Big)^{\frac{1-\alpha\phi}{1-\phi}}\Big(\frac{1-\alpha}{d_t^k}\Big)^{\frac{\phi}{1-\phi}\alpha}$$

$$(5)\ d_t^k = P_t^k - \mathbb{E}_t\Big[\frac{S_{t,t+1}}{1-\theta}\Big((1-\delta_{t+1}) + \Psi'(I_{t+1}/K_t)(I_{t+1}/K_t)^2\Big)P_{t+1}^k\Big]$$

$$(6)\ 1 = P_t^k(1-\Psi(I_t/K_{t-1}) - \Psi'(I_t/K_{t-1})(I_t/K_{t-1}))$$

$$(7)\ PDV_t(\Pi) = \bar{\Pi}_t + \mathbb{E}_t[S_{t,t+1}PDV_{t+1}(\Pi)g_{e,t+1}]$$

$$(8)\ PDV_t(W^f) = \bar{W}_t^f + \mathbb{E}_t[S_{t,t+1}PDV_t(W^f)]$$

$$(9)\ PDV_t(W^+) = (F_{O,s}(\frac{W_t^f}{\bar{w}_t})W_t^f + (1-F_{O,s}(\frac{W_t^f}{\bar{w}_t}))\mathbb{E}_t[O_t^s\bar{w}_t|O_t^s > \frac{W_t^f}{\bar{w}_t}] + \mathbb{E}_t[S_{t,t+1}PDV_{t+1}(W^+)]$$

$$(10)\ Z_t^e = (1-\theta)g_{e,t}Z_{t-1}^e + Z_{0,t}^e$$

$$(11)\ Z_{0,t}^e = \frac{\epsilon\theta\bar{O}_{e,t}^{-(\nu-1)}e^{\mu_t}}{(1-\omega_{r,t}^{-1})(1+\omega_{\ell,t}^{-1})(1-\nu^{-1})(1+\nu\lambda^{-1})} \times \Big(1-(1-\omega_{r,t}^{-1})(\nu-1)(e^{\mu_t}(1-\eta)\bar{O}_{e,t}\frac{PDV_t(W)}{PDV_t(\Pi)})^{-(1+\omega_{\ell,t})}\Big)$$

$$(12)\ 1 = \Big(\frac{r_t}{r_t+\ell_t}\Big(\frac{PDV_t(W)(1-\eta)}{\bar{O}_te^{\mu_t}PDV_t(\Pi)}\Big)^{\ell_t}(1-\phi)^{1-\rho}\frac{(1-\rho)\ell_t^{-1}}{1+(1-\rho)\ell_t^{-1}}$$

$$+\frac{(\bar{O}_t e^{\mu_t} PDV_t(\Pi)/PDV_t(W))^{1-\rho}}{(1-(1-\rho)r_t^{-1})(1+(1-\rho)\ell_t^{-1})}\Big)^{\frac{1}{1-\rho}}$$

$$(13)\ EX_t = \frac{\left(e^{\mu_t}\bar{O}_{e,t}\frac{PDV_t(\Pi)}{(1-\eta)PDV_t(W)}\right)^{-\omega_{\ell,t}}}{(1+\omega_{\ell,t}\omega_{r,t}^{-1})(\omega_{\ell,t}+\nu)}$$

$$(14)\ N_{e,t}^s = (1-\theta)N_{e,t-1}^s + \theta(1-EX_t(1-\eta))\frac{\bar{O}_{e,t}^{-\nu}}{1+\nu\lambda^{-1}}$$

$$(15)\ Y_t = C_t + I_t$$

$$(16)\ Y_t = Z_{e,t}(\bar{\Pi}_t + \bar{N}_t W_t^f) + (1-\epsilon)(1-F_{O,s}(W_t/\bar{w}_t))\mathbb{E}_t[O_t^s\bar{w}_t | O_t^s\bar{w}_t > W_t] + d_t^k K_t$$

$$(17)\ Z_{e,t}\bar{N}_t = \epsilon N_{e,t}^s + (1-\epsilon)F_{O,s}(W_t^f/\bar{w}_t)$$

$$(18)\ W_t = (W_{t-1})^{\kappa}(W_t^f)^{1-\kappa}$$

$$(19)\ \bar{\Pi}_t = \phi^{\frac{\phi}{1-\phi}}\left(\frac{1-\alpha}{W_t}\right)^{\frac{\phi}{1-\phi}(1-\alpha)}\left(\frac{1-\alpha}{d_t^k}\right)^{\frac{\phi}{1-\phi}\alpha}$$

$$(20)\ S_{t,t+1} = (R_{t+1}^x)^{-\rho}\left(\frac{v_{t+1}}{(v_t)^{1/\beta}}\right)^{1-\rho}$$

$$(21)\ R_t^x = \frac{C_t - (1-\beta)(\theta Z_{0,t}^e PDV_t(\bar{\Pi}) + \epsilon(N_{e,t}^s - (1-\theta)N_{e,t-1}^s)PDV_t(W) + (1-\epsilon)\theta PDV_t(W^+))}{\beta(1-\theta)C_{t-1}}$$

$$(22)\ \ln v_t = \beta \ln \mathcal{R}_t[v_{t+1}R_{t+1}^x]$$

C Calibration and additional empirical results

C.1 Derivation of preferences

- Baseline specification: Differential risk adjustment

This section derives the preference specification used in (3.7), replicating the argument in the web appendix of Garleanu and Panageas (2010) and introducing an additional preference-preserving monotone transform $\xi > 0$ for tractability.

Define value in the case of survival or death:

$$\hat{V}_{t+1} = \begin{cases} \tilde{V}_{t+1} & \text{Probability } 1-\theta, 1_S = 1, \\ bA_{t+1} & \text{Probability } \theta, 1_S = 0. \end{cases}$$

Here, bA_{t+1} denotes the utility from bequesting or consuming financial assets $A_{t+1} > 0$.

Consider the specification in Garleanu and Panageas (2010): Agents adjust differentially for different forms of risk. The aversion to survival risk is equal to the inverse of the IES, ψ. In contrast, the aversion to investment risk is ρ:

$$\mathcal{R}[\hat{V}_{t+1}|\mathcal{G}_t] = \mathbb{E}[\mathbb{E}[1_S\tilde{V}_{t+1}^{1-\psi} + (1-1_S)(bA_{t+1})^{1-\psi}|\mathcal{F}_t, \tilde{V}_{t+1}, A_{t+1}]^{\frac{1-\rho}{1-\psi}}|\mathcal{F}_t]^{\frac{1}{1-\rho}}$$

Preference specification (3.7) in the text is the limit point of $b^{1-\psi} = 0$:

$$\begin{aligned}\tilde{V}_t &= \left((1-\tilde{\beta})(\xi C_t)^{1-\psi} + \tilde{\beta}\mathcal{R}_t[\tilde{U}_{t+1}|\mathcal{G}_t]^{1-\psi}\right)^{\frac{1}{1-\psi}} \\ &= \left((1-\beta)(C_t)^{1-\psi} + \beta\mathcal{R}_t[\tilde{U}_{t+1}|\mathcal{G}_t]^{1-\psi}\right)^{\frac{1}{1-\psi}},\end{aligned}$$

Variable	Gross trend growth
Aggregate capital K_t	$\bar{g}_s$
Aggregate investment I_t	$\bar{g}_s$
Aggregate consumption C_t	$\bar{g}_s$
Aggregate entrepreneurial output Y_t	$\bar{g}_s$
Present value wages $PDV_t(W_t)$	$\bar{g}_s$
Present value self-employed income $PDV_t(W_t^+)$	$\bar{g}_s$
Aggregate productivity $Z_{e,t}$	$\bar{g}_s^{\frac{1-\alpha\phi}{1-\phi}} = \bar{g}_e$
Unit labor demand $\bar{N}_t$	$\bar{g}_s^{-\frac{1-\alpha\phi}{1-\phi}} = \bar{g}_e^{-1}$
Unit capital demand $\bar{K}_t$	$\bar{g}_s^{-\frac{\phi(1-\alpha)}{1-\phi}}$
Unit profits $\bar{\pi}_t$	$\bar{g}_s^{-\frac{\phi(1-\alpha)}{1-\phi}}$
Present value profits $PDV_t(G_e\bar{\Pi})$	$\bar{g}_s^{-\frac{\phi(1-\alpha)}{1-\phi}}$
Entering cohort productivity e^{μ_t}	$\bar{g}_s^{\frac{1-\alpha\phi}{1-\phi}} = \bar{g}_e$
Self-employed productivity ω_t	$\bar{g}_s$
Exit rate EX_t	1
Entrepreneurial entry cutoff $\bar{O}_{e,t}$	1
Entrepreneurial labor supply $\bar{N}^s_{e,t}$	1
Self-employed entry cutoff $\bar{O}_{s,t}$	1
Stochastic discount factor $S_{t,t+1}$	1
Return on wealth $R^x_{t,t+1}$	1
Rental rate of capital d^k_t	1
Value of unit wealth v_t	1
Price of capital P^k_t	1

Table B.1: Trend growth rates of variables in model

where $\xi = \left(\frac{1-\tilde{\beta}(1-\theta)}{1-\tilde{\beta}}\right)^{\frac{1}{1-\psi}}$ and $\beta \equiv \tilde{\beta}(1-\theta)$. The re-scaling of the flow utility ensures that the limit of log-utility $\psi \to 1$ is well defined.

- Alternative derivation in terms of uncertainty aversion

Following Hansen and Sargent (2007), define two operators: $\mathbf{T}_I, \mathbf{T}_M$, which incorporate adjustments γ_I, γ_M for investment and mortality risk under robustness:

$$\begin{aligned}\mathbf{T}_I[X] &= -\gamma_I \ln \mathbb{E}[\exp(-\gamma_I^{-1} X)] \\ &= \min_{m(w)\geq 0} \int m(w)\big(X + \gamma_I \ln(m(w))\big) dF(w) \text{ s.t. } \int m(w) dF(w) = 1,\end{aligned}$$

and

$$\begin{aligned}\mathbf{T}_M[X] &= -\gamma_M \ln \mathbb{E}[\exp(-\gamma_M^{-1} X)] \\ &= \min_{m(d)\geq 0} \sum_d p(d) m(d)\big(X + \gamma_M \ln(m(w))\big) \text{ s.t. } \sum_d p(d) m(d) = 1.\end{aligned}$$

Note that, by applying L'Hopital's Rule, as $\gamma \to \infty$, the concern for uncertainty disappears and $\lim_{\gamma\to\infty} \mathbf{T}[X] = \mathbb{E}[X]$.

Similarly to before, define value in the case of survival or death as $\hat{U}_{t+1}$

$$\hat{U}_{t+1} = \begin{cases} \tilde{U}_{t+1} & \text{Probability } 1-\theta, 1_S = 1, \\ bA_{t+1} & \text{Probability } \theta, 1_S = 0. \end{cases}$$

Taking $b \to 0$ and $\gamma_M \to \infty$ therefore implies that:

$$\lim \mathbf{T}_M[\hat{U}_{t+1}|\mathcal{F}_t, \tilde{U}_{t+1}, A_{t+1}]] = (1-\theta)\tilde{U}_{t+1}$$

Now, define the utility recursion in Hansen and Sargent (2007, section 5) using nested robustness operators:

$$\begin{aligned} \tilde{U}_t &= (1-\tilde{\beta})\xi \ln C_t + \tilde{\beta}\mathbf{T}_I[\mathbf{T}_M[\hat{U}_{t+1}|\mathcal{F}_t, \tilde{U}_{t+1}, A_{t+1}]]|\mathcal{F}_t] \\ &\overset{\gamma_M\to\infty, b\to 0}{=} (1-\tilde{\beta})\xi \ln C_t - \tilde{\beta}\gamma_I \ln \mathbb{E}[\exp(-\gamma_I^{-1}(1-\theta)\tilde{U}_{t+1})|\mathcal{F}_t] \end{aligned}$$

Define $\ln V_t \equiv (1-\theta)\tilde{U}_t$ and $\rho = 1 + \gamma_I^{-1}$ and re-arrange to obtain:

$$\begin{aligned} (1-\theta)^{-1} \ln V_t &= (1-\tilde{\beta})\xi \ln C_t - \tilde{\beta}\gamma_I \ln \mathbb{E}[\exp(-\gamma_I^{-1}(1-\theta)\tilde{U}_{t+1})|\mathcal{F}_t] \\ &= (1-\tilde{\beta})\xi \ln C_t - \tilde{\beta}\gamma_I \ln \mathbb{E}[\exp(-\gamma_I^{-1} \ln V_{t+1})|\mathcal{F}_t] \\ &= (1-\tilde{\beta})\xi \ln C_t + \tilde{\beta} \ln\Big(\mathbb{E}[\tilde{V}_{t+1}^{-\gamma_I^{-1}}]^{-\gamma_I}\Big) \\ &= (1-\tilde{\beta})\xi \ln C_t + \tilde{\beta} \ln\Big(\mathcal{R}_{\rho=1+\gamma_I^{-1}}[\tilde{V}_{t+1}]\Big) \\ \Leftrightarrow \ln V_t &= (1-\theta)(1-\tilde{\beta})\xi \ln C_t + (1-\theta)\tilde{\beta} \ln\Big(\mathcal{R}_{\rho=1+\gamma_I^{-1}}[\tilde{V}_{t+1}]\Big) \\ &= (1-\tilde{\beta}(1-\theta)) \ln C_t + (1-\theta)\tilde{\beta} \ln\Big(\mathcal{R}_{\rho=1+\gamma_I^{-1}}[\tilde{V}_{t+1}]\Big), \\ &= (1-\beta) \ln C_t + \beta \ln\Big(\mathcal{R}_{\rho=1+\gamma_I^{-1}}[\tilde{V}_{t+1}]\Big), \end{aligned}$$

for $\xi = \frac{1-\beta(1-\theta)}{(1-\theta)(1-\beta)}$ and using that $\beta \equiv \tilde{\beta}(1-\theta)$.

Thus, the model with uncertainty aversion is observationally equivalent to the model in the body of the paper with log-utility ($\psi = 1$) and investment risk aversion $\rho = 1+\gamma_I^{-1}$.

C.2 Risk sharing without commitment of initial human capital income

Below I argue that, under the assumption of a common stochastic discount factor, the initial lack of commitment by potential entrepreneurs combined with perfect competition by risk-neutral lenders leads to a lack of insurance during the initial stage of occupational choice. Once entrepreneurs can commit, their risk-aversion leads to perfect risk-sharing. Since income derived human capital from human capital is wiped out when agents die, only mortatility risk associated with income from financial capital is insured.

C.2.1 Insurance of human capital income

Assume that at time 0^{-}, a risk-neutral principal commits to a contract to pay $S(z)$ in exchange for the claims to a project paying $z\{G_{i,e,0\to t}\Pi_t\}_{t=0}^{\infty}$ or $(1-\eta)\{W_t\}_{t=0}^{\infty}$. The agent privately observes U_i (while O_i is public information) at time 0 and can decide to seek an outside offer by another principal, or report u to obtain $S(O_i u)$ in exchange for $O_i U_i\{\bar{\Pi}_t\}_{t=0}^{\infty}$. The agent is risk-averse and evaluates different actions according to $\mathcal{R}[X]$, where $X = S(z)$ if she reports z and does not pursue the outside option, or $X = S^o(Z_i)$ for the verified outside option.

Since in equilibrium payoffs are weakly increasing in the expected present discounted value of agents' human capital income, I guess and verify that agents take actions to maximize this present discounted value. Define $\Pi_{i,t}^{+}$ to be this income. Formally:

$$\Pi_t^{+}(Z_{i,0}) \equiv \begin{cases} (1-\eta)W_t & (1-\eta)PDV_0(W) \geq Z_{i,0}PDV_t(G_e\Pi), \\ Z_{i,0}G_{i,e,0\to t}\Pi_t & (1-\eta)PDV_0(W) < Z_{i,0}PDV_t(G_e\Pi), \end{cases}$$

where I define the expected present discounted value of a random payoff $\{X_{t+s}\}_{s=0}^{\infty}$ as $PDV_t(X) = \mathbb{E}[\sum_{s=0}^{\infty} SDF_{t,t+s}X_{t+s}|\mathcal{F}_t]$. Since $G_{i,e,0\to t}$ is *iid* in the cross-section and $O_{i,0}U_{i,0}$ are $\mathcal{F}_0$ measurable, the expected present discounted value of $\Pi_t^{+}(Z_{i,0})$ can be written succinctly as $PDV_0(\Pi_t^{+}(Z_{i,0}))$.

Since agents are risk-averse, it is without loss of generality to focus on payoffs that are a deterministic function of the type only. Let this payoff function at the beginning of the period be defined as $S^{-}: \mathbb{R}^{+} \to \mathbb{R}$, and the end of the period payoff function as $S^{+}: \mathbb{R}^{+} \to \mathbb{R}$.

Note that since the agent is risk-averse, she weakly prefers actuarially fair insurance over noninsurance. The insurance problem can then be solved by backward induction.

Period 0^{+}: Full information, commitment.

(a) Individual Rationality (IR) by financial intermediaries requires that $S^{+}(Z_{i,0}) - PDV(\Pi^{+}(Z_{i,0})) \leq 0$.

(b) Individual Rationality (IR) by agents $S^{+}(Z_{i,0}) - S^{-}(Z_{i,0}) \geq 0$ and $S^{+}(Z_{i,0})$ is preferred to no insurance if the agent accepts the time 0^{+} offer.

Free entry requires that IR by financial intermediaries holds with equality for almost all types $Z_{i,0}$ which accept the time 0^{+} offer.

Since agents are risk averse and $S^{+}(Z_{i,0}) = PDV(\Pi^{+}(Z_{i,0}))$ is strictly preferred to no insurance. Agents' IR requires, therefore, that $PDV(\Pi^{+}(Z_{i,0})) = S^{+}(Z_{i,0}) \geq S^{-}(Z_{i,0})$ for them to be willing to switch.

Period 0^{-}: Limited information, one-sided commitment. Given limited information, S^{-} must be $\mathcal{F}^{-}$ measurable and thus cannot depend on U.

(a) Potential entrepreneurs' IR: $(S^{-}(O_i) =)\mathcal{R}[S^{-}(O_i)|O_i] \geq PDV(W)$ and $S^{-}(O_i)$ is preferred to no insurance.

(b) IC: $S^{-}(O_i) \geq S^{+}(O_i U)$ for all U if type O_i is insured.

(c) Intermediaries' IR: $\mathbb{E}[S^-(O_i) - PDV(\Pi^e(O_iU))|O_i] \leq 0$.

Since S^+ is strictly increasing for $z \geq \underline{z}$, IC implies that $S^-(o) \geq \sup_{u \in \text{supp}(U)} S^+(ou)$. The IC constraint then implies that:

$$\begin{aligned}
&\mathbb{E}[S^-(O_i) - PDV(\Pi^e(O_iU))|O_i] \\
>&\mathbb{E}[S^+(O_iU) - PDV(\Pi^e(O_iU))|O_i] \qquad \text{by IC if } U \text{ is non-degenerate} \\
=&\mathbb{E}[PDV(\Pi^e(O_iU)) - PDV(\Pi^e(O_iU))|O_i] = 0.
\end{aligned}$$

Thus, $S^-(O_i) = \mathbb{E}[S^-(O_i)|O_i] > \mathbb{E}[PDV(\Pi^e(O_iU))|O_i]$. However, this contradicts (IR) for principals.

Since no insurance contract satisfying agents' (IC) and principals' (IR), there is no insurance for idiosyncratic risk at the beginning of the period. However, at the end of the period, actuarial fair insurance is possible.

C.2.2 Insurance of financial capital income

As previously noted, under perfect competition and with commitment, risk-neutral intermediaries offer actuarially fair insurance contracts. Since mortality risk is purely idiosyncratic, the following arrangement, following Blanchard (1985), implements this insurance scheme. An agent holding financial assets A pays A in case of death and otherwise receives a payment a payment of $\theta(1-\theta)^{-1}A$ additional units. Thus, in case of survival, she holds a total of $A\left(1+\theta(1-\theta)^{-1}\right) = A(1-\theta)^{-1}$ units. Intermediaries break even because they receive payments of A from a mass θ of the population and pay $\theta(1-\theta)^{-1}A$ to a mass $1-\theta$ of the population.

C.3 Calibration

C.3.1 Calibrating the skill distribution

The skill distribution is calibrated to match the macro elasticity of labor supply $\frac{d \ln N_t^s}{d \ln W_t}$ and the share of non-entrepreneurs $1 - SE$ in the economy.

Labor supply is given by:

$$N_t^s = \epsilon N_{x,t} + (1-\epsilon)\Phi\Big(\frac{w_t - \mu_w}{\sigma_w}\Big)$$

In steady state: $N^d = N^s = 1 - SE$. Hence:

$$\bar{\Phi} \equiv \Phi\Big(\frac{\bar{w} - \mu_w}{\sigma_w}\Big) = \frac{N^d - \epsilon N_x}{1-\epsilon}$$

Assuming $N_{x,t}$ is constant, the labor supply elasticity is given by:

$$\begin{aligned}\frac{d\ln N_t^s}{d\ln W_t} &\approx \frac{\epsilon\times 0+(1-\epsilon)\phi\Big(\frac{w_t-\mu_w}{\sigma_w}\Big)\frac{w_t}{\sigma_w}}{\epsilon N_{x,t}+(1-\epsilon)\Phi\Big(\frac{w_t-\mu_w}{\sigma_w}\Big)}\\ &\approx (1-\epsilon)\phi\Big(\Phi^{-1}(\bar\Phi)\Big)\frac{\bar w}{\bar N^s\sigma_w}\\ \Leftrightarrow \frac{\sigma_w}{\bar w} &= (1-\epsilon)\phi\Big(\Phi^{-1}(\bar\Phi)\Big)\frac{1}{\bar N^s}\Big(\frac{d\ln N_t^s}{d\ln W_t}\Big)^{-1}\end{aligned}$$

From $\bar\Phi \equiv \Phi\Big(\frac{\bar w-\mu_w}{\sigma_w}\Big)$ it then follows that:

$$\mu_w = \bar w(1-\frac{\sigma_w}{\bar w}\Phi^{-1}(\bar\Phi)) = \bar w - \sigma_w\Phi^{-1}(\bar\Phi))$$

Standard results on the normal distribution imply that the average self-employed income is given by:

$$\mu_w+\sigma_w\frac{\phi\Big(\frac{W_t-\mu_w}{\sigma_w}\Big)}{1-\Phi\Big(\frac{W_t-\mu_w}{\sigma_w}\Big)} = \bar w+\sigma_w\left(\frac{\phi\Big(\frac{W_t-\mu_w}{\sigma_w}\Big)}{1-\Phi\Big(\frac{W_t-\mu_w}{\sigma_w}\Big)}-\Phi^{-1}(\bar\Phi))\right)\equiv \bar w+premium.$$

C.3.2 Implied return to wealth

Given a calibrated value for the ratio of the physical capital stock to human capital income flow as well as the calibrated endowment of young agents, the model yields the steady state return on wealth. Lemma 6 implies that this uniquely defines the return on wealth.

For the computations, I define the data counterpart of human capital income ω as "compensation of employees, paid" plus "Proprietors' income with inventory valuation and capital consumption adjustments" and "Corporate profits with inventory valuation and capital consumption adjustments, domestic industries" from NIPA Table 1.10. I, thereby, exclude what is closer to rental payments on capital, which are a payment to physical rather than human capital in my model. To measure K, I use either the overall stock of private fixed assets plus consumer durables or the consumer durables plus private and governmental fixed assets, all from Table 1.1 in the Fixed Assets Account. From 1979 to 2011, both the median and the mean ratio of $\frac{K}{\omega}$ are 3.7 excluding and 4.6 including government assets. When excluding government compensation of employees from NIPA Table 1.13 and using the private definition of K, the median and mean of $\frac{K}{\omega}$ is 4.3.

C.3.3 Entrepreneurial productivity parameters

Productivity in the entrepreneurial sector is characterized by three parameters: (1) the deterministic observable skill of entrepreneurs μ, (2) the efficiency units of labor embodied in each worker n-scale (after normalizing the steady state wage to unity, i.e. $\bar w = 1$), and (3) the efficiency units of labor embodied in each unit of capital k-scale (since the relative physical amount of capital is pinned down by observables). These can be used to fit the entry rate of entrepreneurs, the ratio of human capital income to the physical capital stock, and the level of labor demand.

Total productivity of entrants is given in equation (B.7) as:

$$
\begin{aligned}
Z_{e,0,t} &= \\
&= \underbrace{\frac{\epsilon\theta\bar{O}_{e,t}^{-\nu}}{(1+\nu\lambda^{-1})}}_{=\theta\times\text{entry}} \frac{e^{\mu_t}\bar{O}_{e,t}}{(1-r_t^{-1})(1+\ell_t^{-1})(1-\nu^{-1})}\Big(1-\frac{1-r_t^{-1}}{1+\ell_t/r_t}\frac{\nu-1}{\ell_t+\nu}\bar{U}_t(\bar{O}_{e,t})^{1+\ell_t}\Big), \\
\bar{U}(\bar{O}) &\equiv (1-\eta)\bar{O}\frac{PDV(\bar{w})}{PDV(\bar{\Pi})e^{\mu}} \overset{\text{in st.st.}}{=} (1-\eta)\bar{O}\frac{\bar{w}}{\bar{\Pi}e^{\mu}}
\end{aligned} \tag{C.1}
$$

Given Lemma 1, $\bar{U}(\bar{O})$ depends only on parameters other than those calibrated in this section. $\bar{Z}_{e,0}$ and hence $\bar{Z}_e = \frac{\bar{Z}_{e,0}}{1-\theta}$ are therefore linear in $e^{\mu}\bar{O}$. Formally, this means that entrepreneurial productivity can be written as:

$$\bar{Z}_e = \hat{Z}_e^c \times e^{\mu}\bar{O}$$

Entry rate is given by (B.1) as:

$$\frac{\epsilon\theta\bar{O}_{e,t}^{-\nu}}{(1+\nu\lambda^{-1})} = \frac{\epsilon\theta(g(\ell,r)e^{-\mu}\frac{\bar{w}}{\bar{\Pi}})^{-\nu}}{(1+\nu\lambda^{-1})},$$

which can be set to equal its empirical counterpart in steady sate and is a function of μ, n-scale, and k-scale. To keep $\bar{O}$ and therefore entry constant, it is necessary and sufficient to keep $e^{\mu}\frac{\bar{\Pi}}{\bar{w}}$ constant, or equivalently:

$$\frac{e^{\mu}\bar{\Pi}}{\bar{w}} \propto e^{\mu_{new}}(\text{n-scale}_{new})^{\frac{\phi-\alpha\phi}{1-\phi}}(\text{k-scale}_{new})^{\frac{\phi\alpha}{1-\phi}} = e^{\mu_{old}}(\text{n-scale}_{old})^{\frac{\phi-\alpha\phi}{1-\phi}}$$

Unit factor demand for capital and labor is given by:

$$
\begin{aligned}
\bar{k} &= \phi^{\frac{1}{1-\phi}}\left(\frac{1-\alpha}{\bar{w}/\text{n-scale}}\right)^{\frac{(1-\alpha)\phi}{1-\phi}}\left(\frac{\alpha}{\bar{\bar{d}}}\right)^{\frac{1-(1-\alpha)\phi}{1-\phi}}\text{k-scale}^{\frac{1-(1-\alpha)\phi}{1-\phi}} \\
&= \phi^{\frac{1}{1-\phi}}\left(\frac{1-\alpha}{\bar{w}}\right)^{\frac{(1-\alpha)\phi}{1-\phi}}\left(\frac{\alpha}{\bar{\bar{d}}}\right)^{\frac{1-(1-\alpha)\phi}{1-\phi}}\left(\text{n-scale}^{(1-\alpha)\phi}\text{k-scale}^{1-(1-\alpha)\phi}\right)^{\frac{1}{1-\phi}} \\
&\equiv \hat{\kappa}^c\Big(\text{n-scale}^{(1-\alpha)\phi}\text{k-scale}^{1-(1-\alpha)\phi}\Big)^{\frac{1}{1-\phi}} \\
\bar{N} &= \phi^{\frac{1}{1-\phi}}\left(\frac{1-\alpha}{\bar{w}}\right)^{\frac{1-\alpha\phi}{1-\phi}}\left(\frac{\alpha}{\bar{r}}\right)^{\frac{\alpha\phi}{1-\phi}}\text{n-scale}^{\frac{1-\alpha\phi}{1-\phi}}\text{k-scale}^{\frac{\phi\alpha}{1-\phi}} \\
&\equiv \hat{N}^c \times \text{n-scale}^{\frac{1-\alpha\phi}{1-\phi}}\text{k-scale}^{\frac{\phi\alpha}{1-\phi}}
\end{aligned}
$$

The following algorithm describes how units for capital input, labor input, and entrepreneurial span of control can be chosen to match the entry rate of entrepreneurs, the ratio of human capital income to the physical capital stock, and the level of labor demand.

1. Set k-scale$^{(0)} = 1$ and solve $1 = \max\{(1-\eta), \mathbf{o}\frac{\hat{\Pi}}{\bar{w}}\mathcal{R}[U]\}$ for $\mathbf{o} = e^{\mu}\bar{O}\text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\text{k-scale}^{\frac{\phi\alpha}{1-\phi}}$.

2. Choose n-scale$^{(0)}$ such that given y

$$\hat{Z}^c \times e^{\mu}\bar{O}\text{n-scale}^{\frac{1-\alpha\phi}{1-\phi}}\text{k-scale}^{\frac{\phi\alpha}{1-\phi}}\hat{N}^c = \hat{Z}^c \times \mathbf{o} \times \text{n-scale}^{(0)}\hat{N}^c \stackrel{!}{=} \bar{N}^{target}.$$

3. Choose $\mu^{(1)}$ to guarantee the targeted entry level:

$$\frac{\epsilon\theta\left(\bar{O}\right)^{-\nu}}{(1+\nu\lambda^{-1})} = \frac{\epsilon\theta\left(\frac{\mathbf{o}}{e^{\mu}\text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\text{k-scale}^{\frac{\alpha\phi}{1-\phi}}}\right)^{-\nu}}{(1+\nu\lambda^{-1})} \stackrel{!}{=} E^{target}$$

Thus

$$e^{\mu^{(1)}} \propto \frac{\mathbf{o}}{\text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\text{k-scale}^{\frac{\alpha\phi}{1-\phi}}},$$

where $1-\alpha$ is the labor share.

4. Choose k-scale$^{(1)}$ according to:

$$\begin{aligned}
x &\equiv \frac{\omega}{K} + \frac{\text{net end}}{K} \\
&= \left(\frac{1-\phi}{\phi}\frac{\bar{d}}{\alpha\text{k-scale}^{(1)}} + \frac{\bar{w}(1-SE\epsilon-(1-\epsilon)Entry)+\epsilon SE(\bar{w}+premium)}{\hat{k}e^{\mu^{(1)}}\hat{Z}^c\bar{O}\left(\text{n-scale}^{(1)(1-\alpha)\phi}\text{k-scale}^{1-(1-\alpha)\phi}\right)^{\frac{1}{1-\phi}}}\right) \\
&= \left(\frac{1-\phi}{\phi}\frac{\bar{d}}{\alpha\text{k-scale}^{(1)}} + \frac{\bar{w}(1-SE\epsilon-(1-\epsilon)Entry)+\epsilon SE(\bar{w}+premium)}{\hat{k}e^{\mu^{(1)}}\hat{Z}^c\bar{O}\left(\text{n-scale}^{(1)(1-\alpha)\phi}\text{k-scale}^{\alpha\phi}\right)^{\frac{1}{1-\phi}}\text{k-scale}}\right) \\
&= \left(\frac{1-\phi}{\phi}\frac{\bar{d}}{\alpha\text{k-scale}^{(1)}} + \frac{\bar{w}(1-SE\epsilon-(1-\epsilon)Entry)+\epsilon SE(\bar{w}+premium)}{\hat{k}\mathbf{o}\hat{Z}^c\bar{O}\text{k-scale}^{(1)}}\right),
\end{aligned}$$

where ω is the total human capital income, i.e.:

$$\omega = Z\bar{\Pi} + \bar{w}(1-SE\epsilon-(1-\epsilon)Entry)+\epsilon SE(\bar{w}+premium),$$

and using that unit profits can be written as:

$$\begin{aligned}
\bar{\Pi} &= (1-\phi)\phi^{\frac{\phi}{1-\phi}}\left(\frac{1-\alpha}{\bar{w}}\right)^{\frac{\phi(1-\alpha)}{1-\phi}}\left(\frac{\alpha}{\bar{r}}\right)^{\frac{\phi\alpha}{1-\phi}}\text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\text{k-scale}^{\frac{\phi\alpha}{1-\phi}} \\
&\equiv \hat{\Pi} \times \text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\text{k-scale}^{\frac{\phi\alpha}{1-\phi}}
\end{aligned}$$

5. Compute the entry cutoff $\bar{O}$:

$$\bar{O} = \frac{\mathbf{o}}{e^{\mu}\text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\,\text{k-scale}^{\frac{\phi\alpha}{1-\phi}}}$$

6. Verify that $\bar{O} > 1$ and $\bar{U} > \bar{O}$, i.e.

$$\frac{\bar{w}}{\bar{\Pi}e^{\mu}}(1-\eta) = \frac{\bar{w}}{\hat{\Pi}e^{\mu} \times \text{n-scale}^{\frac{\phi(1-\alpha)}{1-\phi}}\,\text{k-scale}^{\frac{\phi\alpha}{1-\phi}}}(1-\eta) > \bar{O}$$

C.4 Impulse-response functions

C.4.1 IRFs to Independent Shocks

The IRFs are computed at the posterior mean for the establishment-level estimates in Table 3 and the calibration in Table 2. The exception is the stochastic process for the depreciation shock, which is turned off for the estimation. I calibrated the AR(1) process with a persistence of 0.55 and a standard deviation of 0.1.

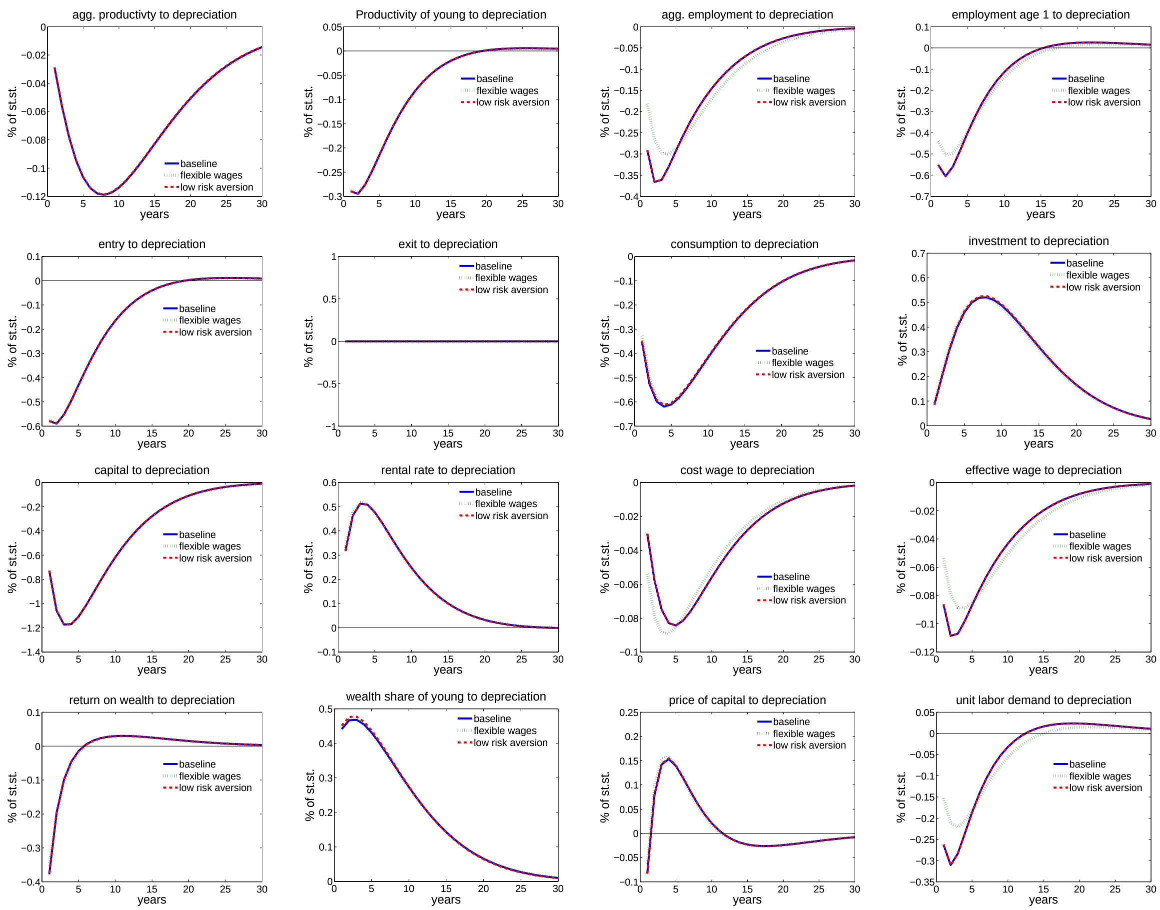


Figure C.1: Impulse response to one standard deviation depreciation shock in baseline log-linear model at posterior mean

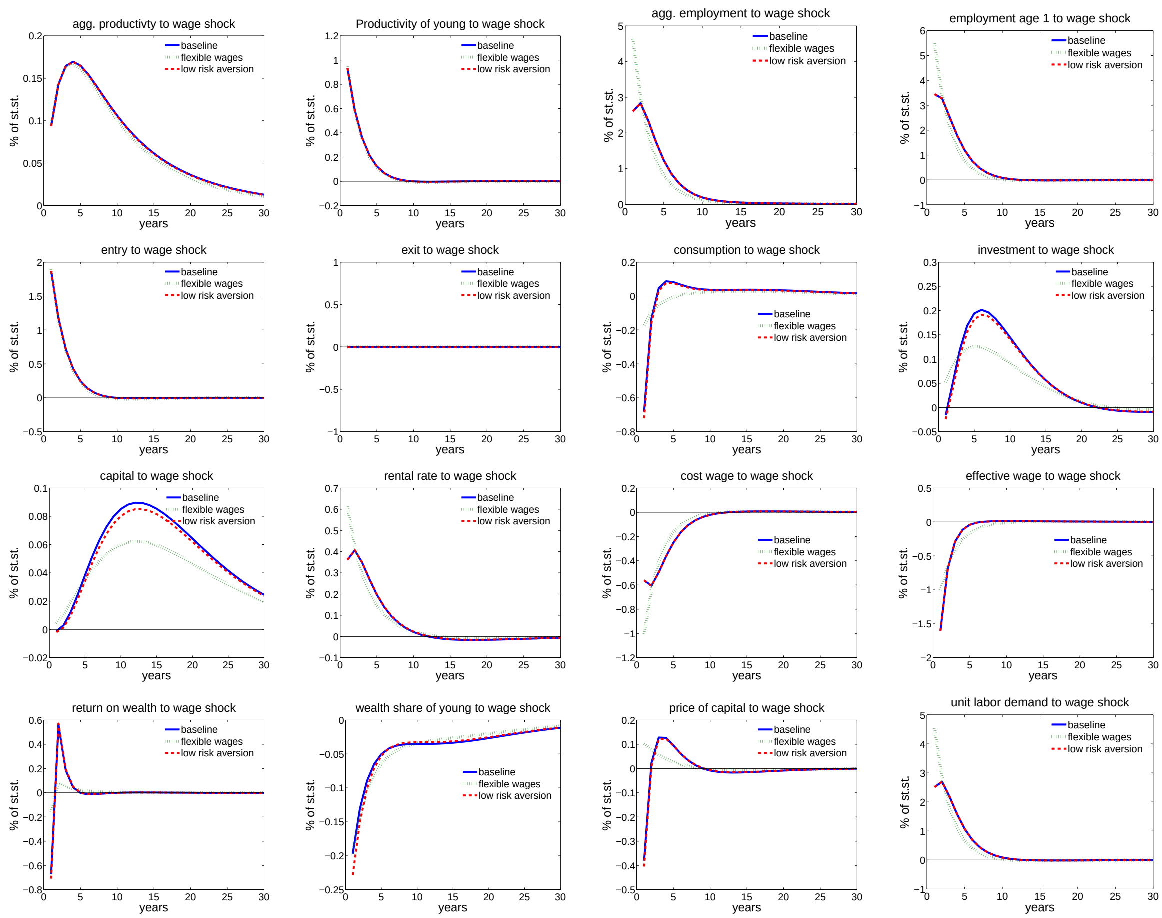


Figure C.2: Impulse response to one standard deviation wage shock in baseline log-linear model at posterior mean

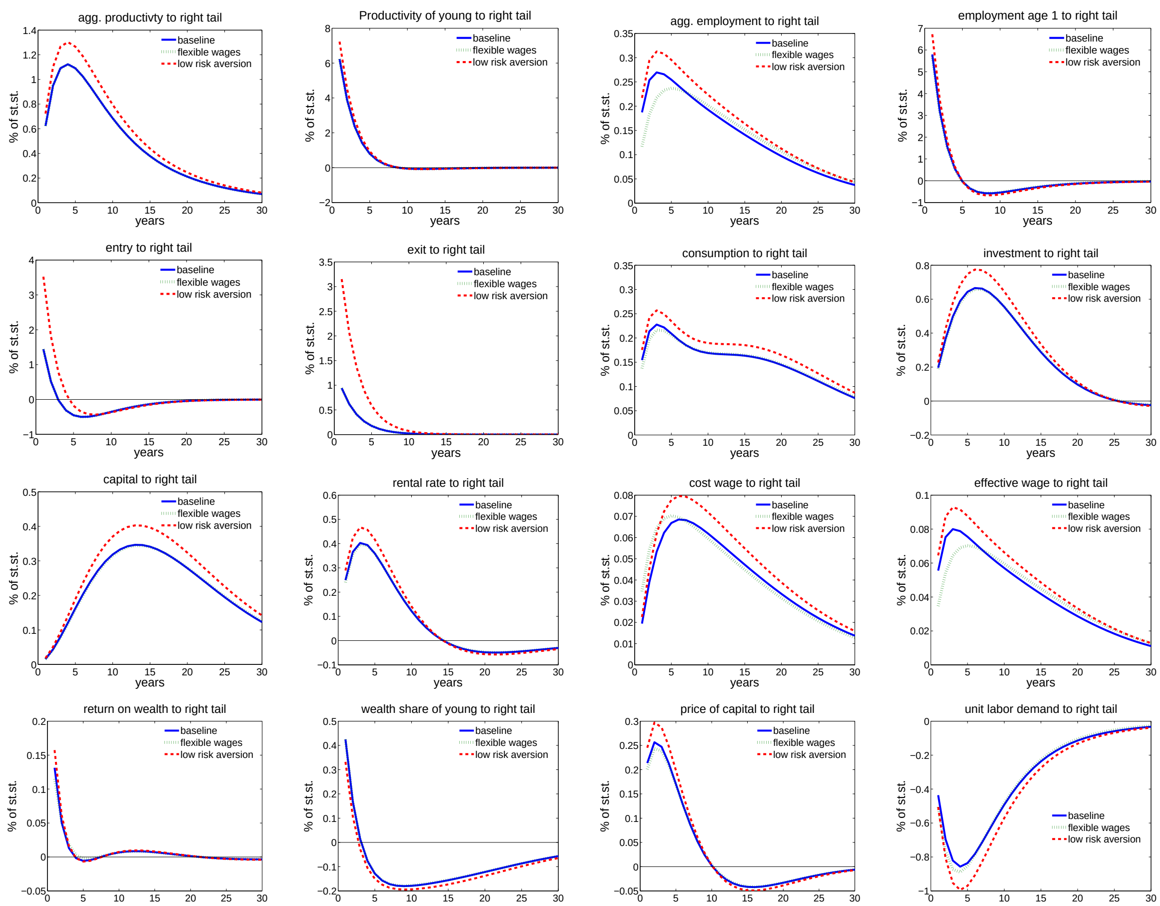


Figure C.3: Impulse response to one standard deviation upside risk in baseline log-linear model at posterior mean

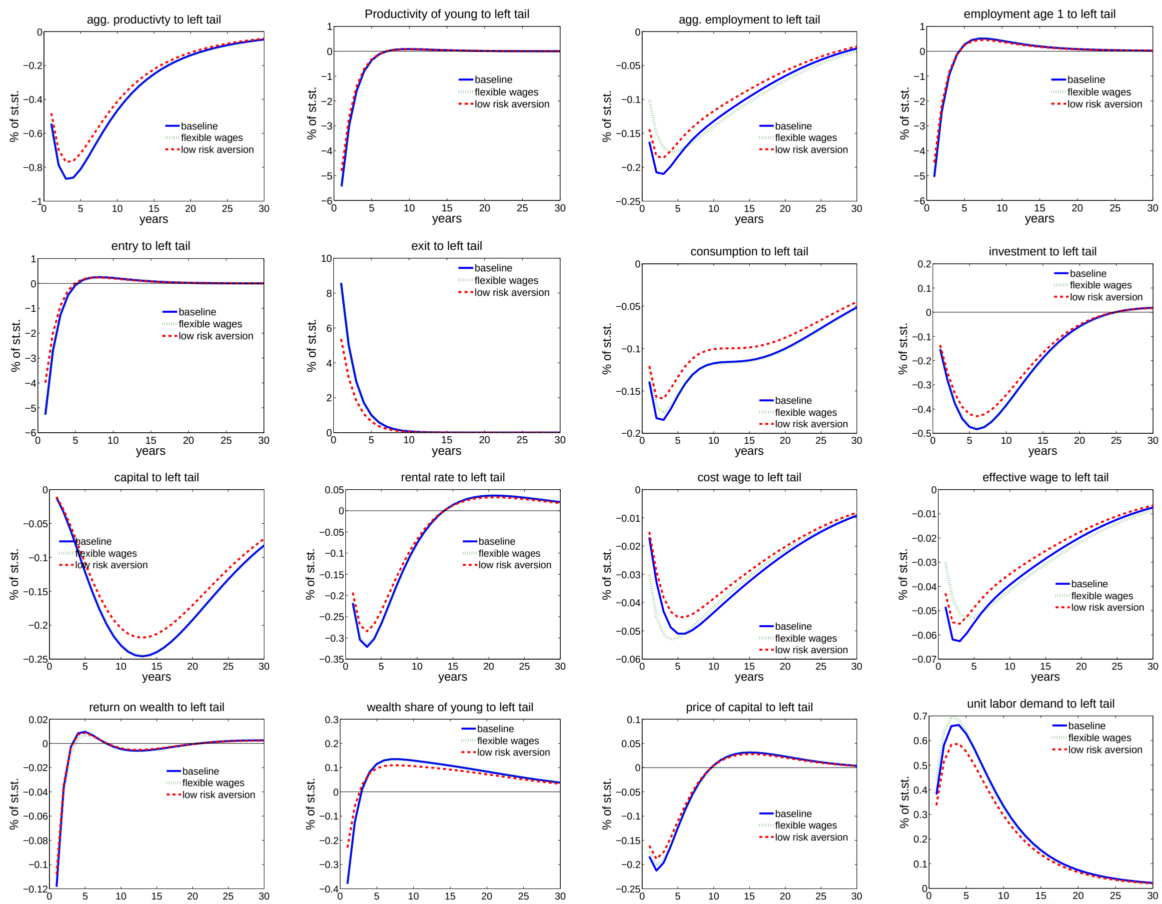


Figure C.4: Impulse response to one standard deviation downside risk shock in baseline log-linear model at posterior mean

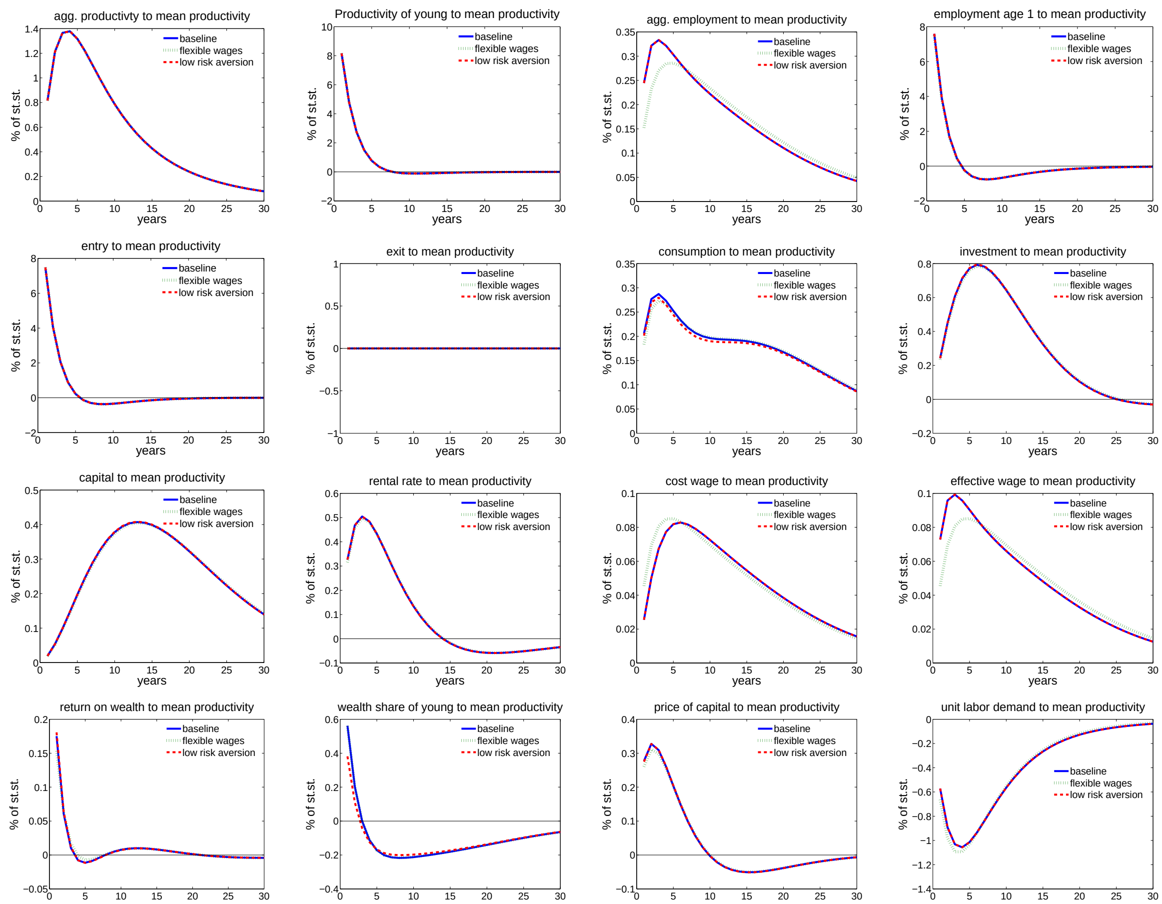


Figure C.5: Impulse response to one standard deviation mean productivity shock in baseline log-linear model at posterior mean

C.4.2 IRFs to Mean Preserving Spread

Notice that the expected value of the unexpected productivity component is given by:

$$\mathbb{E}_t[U_{it}] = \int_{-\infty}^{\infty} \frac{\ell_t r_t}{\ell_t + r_t} \exp(u + \min\{\ell_t u, -r_t u\}) du = \frac{\ell_t}{1+\ell_t}\frac{r_t}{r_t - 1} = \frac{1}{1+\ell_t^{-1}}\frac{1}{1-r_t^{-1}}$$

The mean of $\mathbb{E}_t[U_{it}e^{\mu_t}]$ is consequently given by $\frac{1}{1+\ell_t^{-1}}\frac{1}{1-r_t^{-1}}e^{\mu_t}$.

In practice, I specify the shock process as

$$\begin{aligned}
\ln \ell_t &= (1-\rho_\ell)\ln \bar{\ell} + \rho_\ell \ln \ell_{t-1} - \sigma_\ell \epsilon_t^\ell \\
\ln(r_t - 1) &= (1-\rho_r)\ln(\bar{r}-1) + \rho_r \ln(r_{t-1}-1) - \sigma_r \epsilon_t^r \\
\ln \mu_t &= (1-\rho_\mu)\ln(\bar{\mu}-1) + \rho_\mu \ln \mu_{t-1} + \sigma_\mu \epsilon_t^\mu,
\end{aligned}$$

where the different signs of the shocks ensure that tail risk shocks increase the tail risk.

Re-write the mean accordingly:

$$\begin{aligned}
\mathbb{E}_t[U_{it}e^{\mu_t}] &= \frac{1}{1+\ell_t^{-1}}\frac{1}{1-r_t^{-1}}e^{\mu_t} \\
&= \frac{1}{1+\exp(-\ln \ell_t)}\frac{1}{1-(1+\exp(\ln(r_t-1)))^{-1}}e^{\mu_t} \\
&= \frac{1}{1+\exp(-\ln \ell_t)}(1+\exp(-\ln(r_t-1)))e^{\mu_t}
\end{aligned}$$

The total derivative is given by:

$$d\mathbb{E}_t[U_{it}e^{\mu_t}] = \mathbb{E}_t[U_{it}e^{\mu_t}]\left(\frac{1}{1+\exp(-\ln \ell_t)}d\ln \ell_t - \frac{1}{1+\exp(-\ln(r_t-1))}d\ln(r_t-1) + d\mu_t\right)$$

Thus, the mean preserving spread requires that:

$$d\mu_t = -\frac{1}{1+\exp(-\ln \bar{\ell})}d\ln \ell_t + \frac{1}{1+\exp(-\ln(\bar{r}-1))}d\ln(r_t-1)$$

Note that to compute the IRF, a common persistence has to be imposed. In the computations, I impose for clarity that $\rho_\circ = 0$.

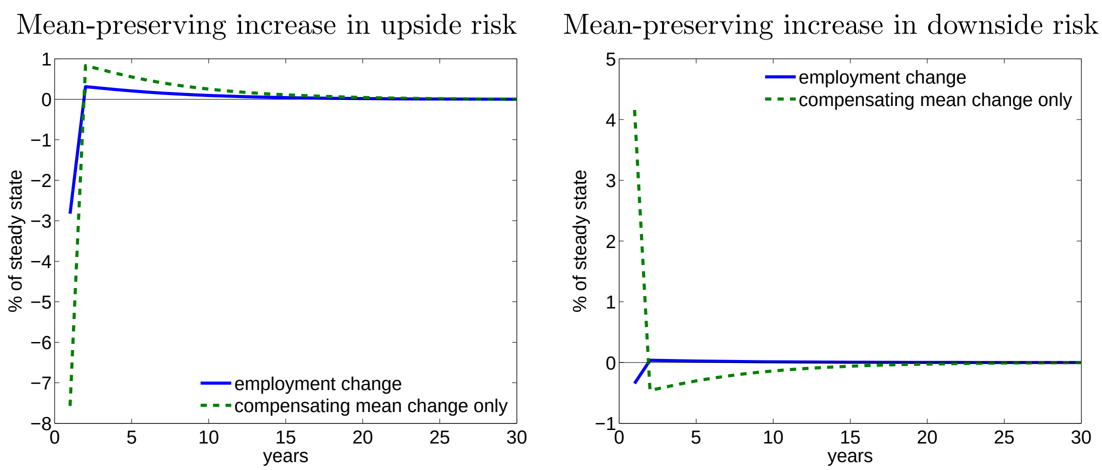


Figure C.6: Employment response to mean-preserving spread upside and downside risk shocks in baseline log-linear model at posterior mean

C.5 Entry and exit with borrowing and lending

In the baseline model, the fixed cost of entry consists entirely of a loss of wages equal to a fraction η of lifetime wages. To generalize, a pecuniary entry cost can be introduced as $\zeta\Phi\bar{\omega}$, where $\zeta \in (0,1)$ is a scaling variable, $\Phi > 0$ is a fixed cost parameter, and $\bar{\omega}$ is the flow of human capital income in stationary equilibrium. To limit the number of free parameters, I assume that households are endowed with $\zeta\bar{\omega}$ at birth. They can finance a fraction γ of the entry cost with risky debt: They default on their debt obligations when they exit. Assuming there is a pooling equilibrium in the loan market, the gross intra-period interest rate is $(1-\overline{EX})^{-1}$, resulting in endogenous propagation. In terms of the calibrated model above, the aggregate endowment net of entry cost is therefore net end.$_t = \theta(1 - \text{Entry}_t)\zeta\bar{\omega}$.

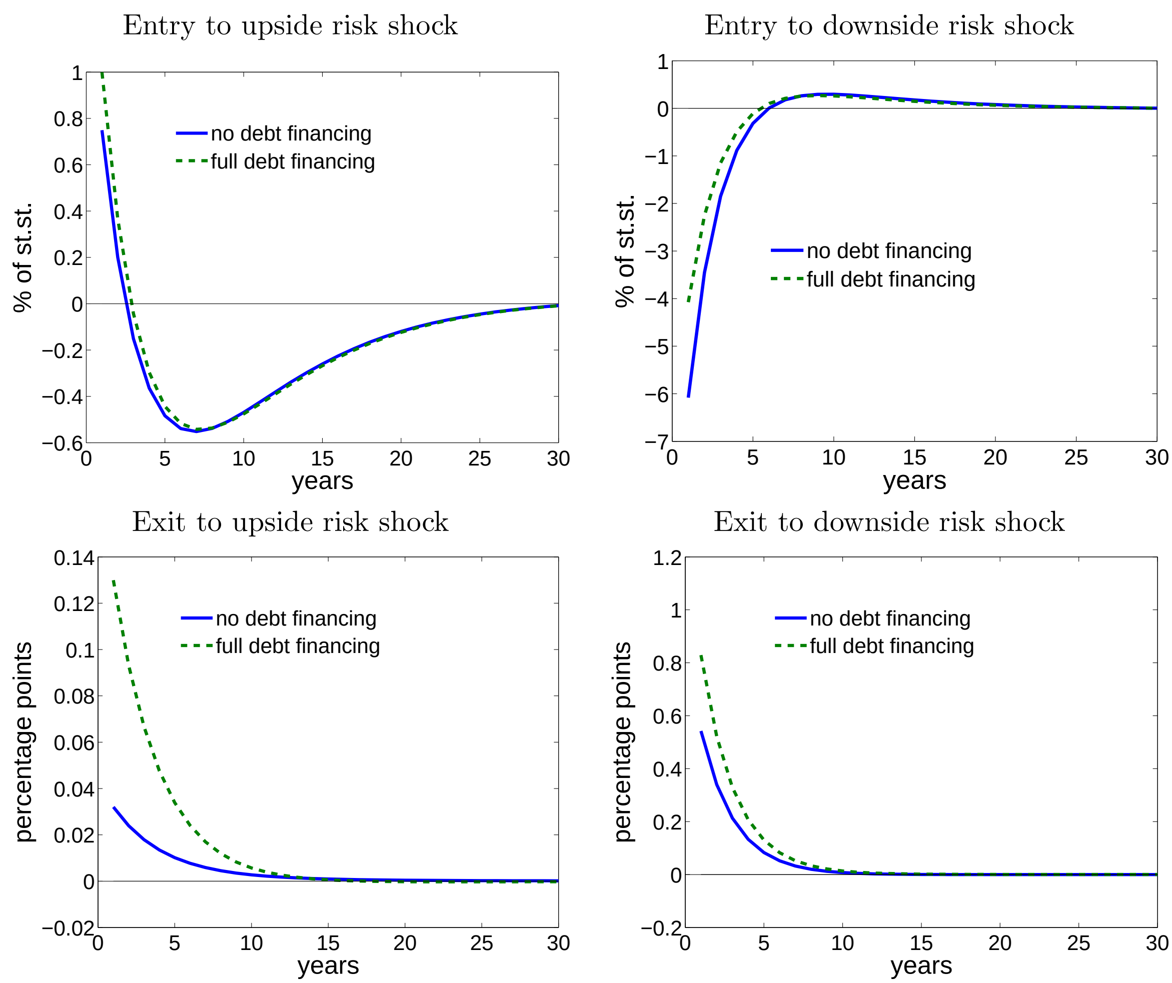
It can be shown that in this case the entry criterion becomes:

$$PDV(W) + \zeta K = \mathcal{R}[\max\{(1-\eta)PDV(W) + \zeta\frac{(1-\gamma)\Phi\bar{\omega}}{1-\text{EX}}, O \times UPDV(\Pi)\} + \zeta\bar{\omega}].$$

The exit criterion becomes: $(1-\eta)PDV(W) > O \times UPDV(\Pi) - \zeta\frac{\gamma\Phi\bar{\omega}}{1-\text{EX}}$. The exit rate now solves a fixed point problem, reflecting the fact that higher exit rates make default more and entry less attractive. Otherwise the solution to the economy is unchanged.

Figure C.7 provides the IRF for two numerical examples, using $\zeta = 1, \gamma \in \{0,1\}, \Phi = 1$. With $\gamma = 0$, there is no loan market, while with $\gamma = 1$, the entrepreneur finances the entire fixed cost with external debt. With full external financing, the pecuniary entry cost is paid only if the entrepreneur is successful, providing partial insurance against failure. Thus entry becomes even more attractive when upside risk increases. At the same time, with the full external financing of the pecuniary entry cost, exit rates rise as agents select less carefully and have a larger incentive to exit when downside risk increases.

Figure C.7: Elasticity of entry, exit, and productivity with respect to risk with fixed cost



C.6 Data

I use data on employment from the Business Dynamics Statistics (BDS). The BDS are based on the Longitudinal Business Database (LBD), which covers all employer business in the US since 1976 and is based on administrative data. The BDS is organized by sector, firm or establishment age, and firm or establishment employment size. Besides employment data, it also contains data on the number of entering and exiting businesses.

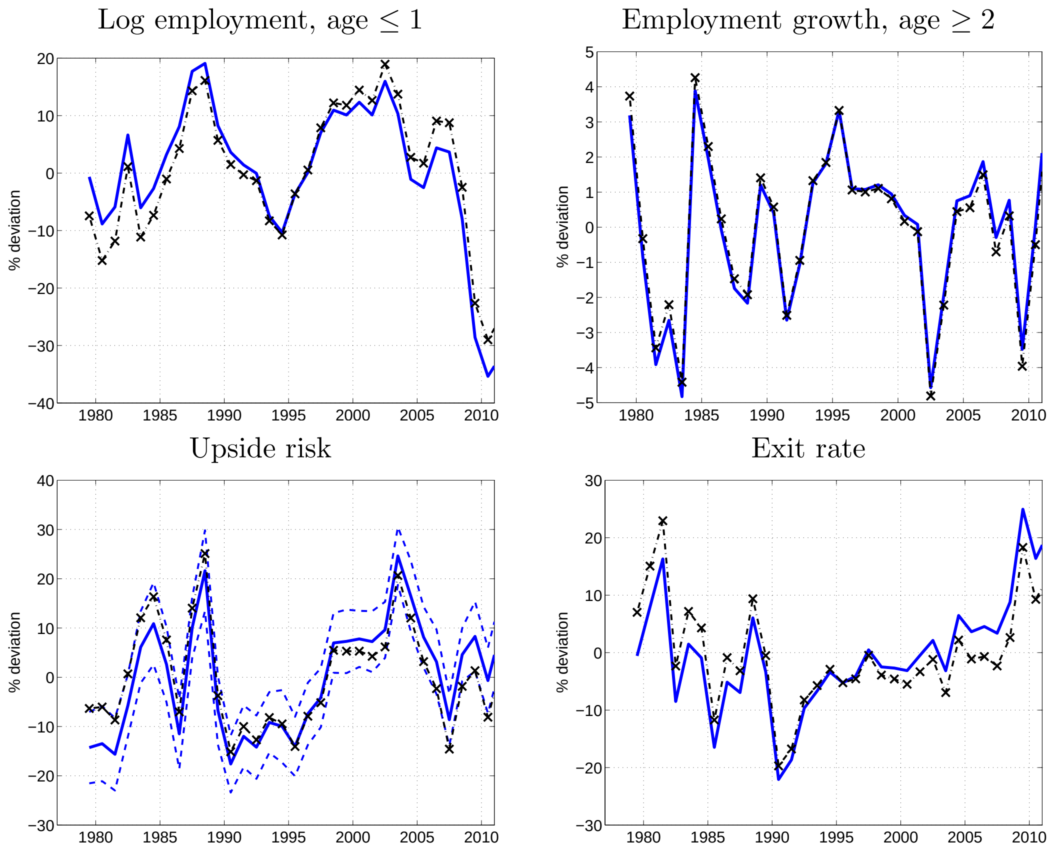


Figure C.8: Data used in estimation: Establishment-level

Note: Blue solid lines display the data without detrending used in the baseline estimation. Black dashed-crossed lines display data after trend removal.

Figure C.9: Data used in estimation: Firm-level

Note: Blue solid lines display the data without detrending. Black dashed-crossed lines display data after trend removal. In the baseline estimation, only the upside risk time series is detrended.

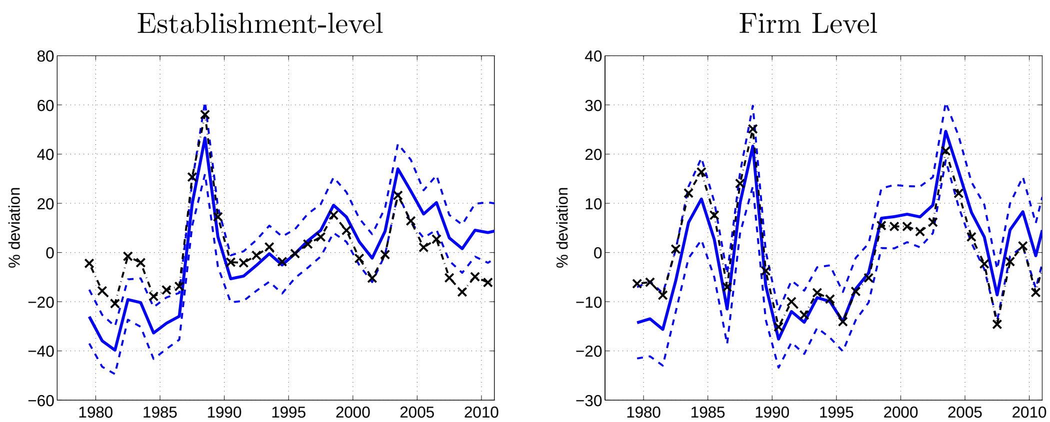


Figure C.10: Upside risk measured at 20+ employees: Establishment and firm level data

Note: Blue solid lines display the data without detrending. Black dashed-crossed lines display data after trend removal.

C.7 Comparison of datasets

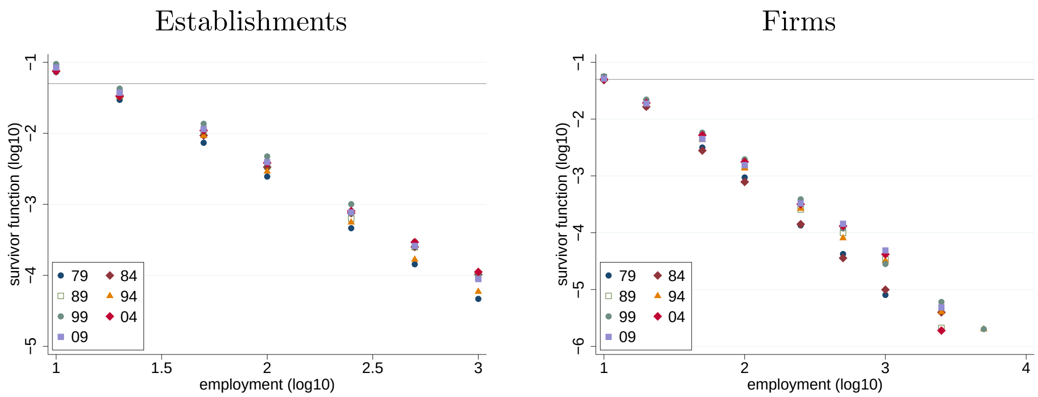


Figure C.11: Log countercumulative distribution function vs. log size for businesses age 0 at five-year intervals

Figure C.12: Young businesses: net job creation and employment share by type of business, 1979–2010

C.7.1 Reduced form VAR evidence on risk and employment, average size, and entry

Table C.1: Impulse-response on impact to upside risk and exit rate shocks: Aggregate results

Specification	New bus. employment	Entry	Avg size
	Establishments, upside risk		
Baseline	0.91(0.10,1.75)	0.49(-0.10,1.11)	0.41(-0.40,1.25)
Baseline, cutoff: 20+	0.78(-0.03,1.62)	-0.04(-0.60,0.52)	0.82(0.00,1.67)
Age 0	3.08(1.55,4.71)	1.18(0.20,2.18)	1.91(1.16,2.71)
Age 1	0.49(-0.27,1.29)	-1.05(-1.69,-0.44)	1.54(0.78,2.36)
	Establishments, exit rate		
Baseline	-1.42(-2.22,-0.67)	-0.39(-0.98,0.19)	-1.03(-1.84,-0.26)
Baseline, cutoff: 20+	-1.11(-1.91,-0.36)	-0.33(-0.88,0.20)	-0.78(-1.58,-0.01)
Age 0	-3.60(-5.07,-2.26)	-2.32(-3.26,-1.46)	-1.27(-1.98,-0.62)
Age 1	-0.75(-1.51,-0.02)	-0.41(-1.00,0.16)	-0.34(-1.09,0.38)
	Firms, upside risk		
Baseline	2.94(2.25,3.73)	1.89(1.40,2.44)	1.06(0.41,1.73)
Baseline, cutoff: 20+	2.97(2.26,3.78)	1.19(0.71,1.70)	1.78(1.15,2.48)
Age 0	3.29(2.08,4.62)	0.77(0.05,1.53)	2.51(1.78,3.35)
Age 1	1.78(0.99,2.64)	0.13(-0.40,0.66)	1.65(1.03,2.34)
	Firms, exit rate		
Baseline	0.03(-0.56,0.63)	-0.49(-0.93,-0.08)	0.52(-0.09,1.16)
Baseline, cutoff: 20+	-0.67(-1.29,-0.07)	-0.36(-0.82,0.08)	-0.31(-0.91,0.28)
Age 0	-3.23(-4.36,-2.20)	-2.11(-2.80,-1.49)	-1.12(-1.80,-0.46)
Age 1	-0.80(-1.57,-0.06)	-0.97(-1.49,-0.49)	0.17(-0.42,0.76)

Shown are the posterior median and the 68% credible set in parentheses for different specifications of new businesses and the corresponding risk measures.

Table C.2: Impulse-response on impact to upside risk and exit rate shocks: Sectoral results

SIC code	7	10	15-19	20-39	40-49	50-51	52-59	60-69	70-89
	Establishments, upside risk								
Employment	2.5	-1.3	-2.7	0.2	2.6	1.1	-0.4	4.1	2.4
to upside	(1.0,4.0)	(-4.6,2.0)	(-4.4,-1.0)	(-1.3,1.7)	(0.6,4.7)	(0.1,2.2)	(-1.3,0.5)	(2.6,5.8)	(1.1,3.7)
Entry	-2.2	-2.1	-1.2	-1.1	-0.7	-1.4	-1.1	0.4	0.5
to upside	(-2.9,-1.6)	(-3.3,-0.9)	(-2.1,-0.4)	(-1.7,-0.4)	(-1.6,0.2)	(-2.3,-0.5)	(-1.7,-0.6)	(-0.7,1.5)	(-0.3,1.3)
Avg size	4.7	0.8	-1.4	1.2	3.3	2.5	0.8	3.8	1.9
to upside	(3.4,6.2)	(-1.9,3.5)	(-2.6,-0.3)	(-0.0,2.5)	(1.4,5.4)	(1.6,3.4)	(0.0,1.5)	(1.7,5.9)	(0.7,3.1)
	Establishments, exit rate								
Employment	-3.2	-13.3	-6.1	-0.7	-1.7	-2.6	-1.0	-0.7	-0.9
to exit	(-4.6,-1.9)	(-16.1,-10.9)	(-7.6,-4.8)	(-2.2,0.7)	(-3.6,0.3)	(-3.6,-1.7)	(-1.9,-0.2)	(-2.1,0.8)	(-2.1,0.3)
Entry	-0.7	-3.8	-1.0	-1.0	-0.7	-0.7	-0.2	1.5	-2.0
to exit	(-1.3,-0.2)	(-4.9,-2.9)	(-1.8,-0.2)	(-1.6,-0.4)	(-1.5,0.2)	(-1.6,0.1)	(-0.7,0.3)	(0.5,2.5)	(-2.8,-1.4)
Avg size	-2.5	-9.4	-5.1	0.3	-1.0	-1.9	-0.8	-2.1	1.2
to exit	(-3.7,-1.4)	(-11.9,-7.3)	(-6.1,-4.3)	(-0.9,1.5)	(-2.9,0.9)	(-2.6,-1.2)	(-1.5,-0.1)	(-4.1,-0.2)	(0.1,2.3)
Emp. share	0.9	0.7	5.5	9.6	5.8	5.8	27.7	8.1	36.0
	Firms, upside risk								
Employment	-1.2	4.5	-2.8	1.3	-0.2	0.7	-0.5	-0.5	2.4
to upside	(-2.5,0.1)	(0.2,8.9)	(-4.1,-1.6)	(0.0,2.6)	(-1.3,1.0)	(-0.4,1.8)	(-1.2,0.1)	(-1.6,0.5)	(1.4,3.6)
Entry	-1.8	-0.9	-1.4	-0.2	-0.9	-0.1	-0.7	0.1	-0.7
to upside	(-2.5,-1.2)	(-2.1,0.2)	(-2.0,-0.9)	(-0.7,0.3)	(-1.7,-0.2)	(-0.7,0.4)	(-1.0,-0.5)	(-0.5,0.8)	(-1.5,0.1)
Avg size	0.6	5.4	-1.4	1.4	0.8	0.8	0.2	-0.7	3.2
to upside	(-0.8,2.1)	(1.9,9.2)	(-2.5,-0.4)	(0.3,2.6)	(-0.2,1.7)	(-0.1,1.7)	(-0.4,0.8)	(-1.5,0.1)	(2.2,4.3)
	Firms, exit rate								
Employment	-2.7	-10.1	-4.0	-1.5	-1.0	-1.6	-1.1	-2.5	-2.5
to exit	(-4.0,-1.6)	(-14.1,-6.4)	(-5.1,-3.1)	(-2.7,-0.4)	(-2.2,0.1)	(-2.7,-0.7)	(-1.8,-0.5)	(-3.6,-1.6)	(-3.5,-1.6)
Entry	-0.1	-1.2	-0.2	-1.5	0.2	-1.8	-0.1	-0.7	-2.6
to exit	(-0.7,0.5)	(-2.3,-0.1)	(-0.7,0.3)	(-2.0,-1.1)	(-0.5,0.9)	(-2.3,-1.3)	(-0.3,0.2)	(-1.3,-0.1)	(-3.3,-1.9)
Avg size	-2.6	-8.9	-3.9	-0.0	-1.2	0.1	-1.1	-1.9	0.1
to exit	(-4.0,-1.3)	(-12.2,-5.9)	(-4.8,-3.1)	(-1.1,1.1)	(-2.1,-0.3)	(-0.7,1.0)	(-1.6,-0.6)	(-2.6,-1.2)	(-0.9,1.0)
Emp. share	1.4	0.7	8.6	8.4	4.3	5.2	26.1	5.5	40.2

Shown are the posterior median and the 68% credible set in parentheses. New business employment, exit, and risk are defined for businesses up to age 1, as in the baseline aggregate specification. SIC classifications underlying the sectors: 7 (agriculture), 10 (mining), 15-19 (construction), 20-39 (manufacturing), 40-49 (transport and utilities), 50-51 (wholesale), 52-59 (retail), 60-69 (finance, insurance, real estate), 70-89 (business and other services).

C.7.2 Historical variance decomposition by subsample

The following tables extend the historical variance decomposition from Table 4 separately for the first and the second half of my sample.

Table C.3: Details of historical variance decomposition: Establishments, 1979–2010

		(1) Upside	(2) Downside	(1) & (2)	other
		Full sample			
Employment, age 1	rel var	0.17 (0.15,0.19)	0.22 (0.15,0.29)	0.38 (0.31,0.47)	1.00 (1.00,1.00)
	corr	0.38 (0.33,0.44)	0.57 (0.54,0.62)	0.72 (0.65,0.76)	1.00 (1.00,1.00)
Employment growth, age 2+	rel var	0.04 (0.03,0.05)	-0.01 (-0.01,0.00)	0.03 (0.02,0.04)	1.00 (1.00,1.00)
	corr	0.20 (0.15,0.26)	-0.04 (-0.07,0.00)	0.14 (0.09,0.19)	1.00 (1.00,1.00)
Exit, age 1	rel var	0.04 (0.02,0.08)	0.96 (0.92,0.98)	1.00 (0.99,1.00)	1.00 (1.00,1.00)
	corr	0.36 (0.32,0.39)	0.99 (0.97,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.03 (0.01,0.04)	0.26 (0.17,0.38)	0.29 (0.19,0.41)	0.81 (0.72,0.91)
	corr	0.19 (0.07,0.34)	0.55 (0.54,0.56)	0.57 (0.54,0.60)	0.80 (0.76,0.83)
Average size	rel var	0.34 (0.30,0.38)	-0.02 (-0.04,0.00)	0.32 (0.28,0.36)	0.33 (0.23,0.45)
	corr	0.58 (0.53,0.61)	-0.19 (-0.38,0.05)	0.53 (0.45,0.59)	0.48 (0.33,0.59)
		1st half: 1979–1994			
Employment, age 1	rel var	0.33 (0.29,0.37)	0.18 (0.13,0.23)	0.50 (0.43,0.58)	1.00 (1.00,1.00)
	corr	0.49 (0.44,0.54)	0.31 (0.27,0.37)	0.63 (0.59,0.68)	1.00 (1.00,1.00)
Employment growth, age 2+	rel var	0.03 (0.02,0.04)	-0.02 (-0.04,-0.01)	0.00 (-0.01,0.02)	1.00 (1.00,1.00)
	corr	0.15 (0.11,0.20)	-0.18 (-0.20,-0.15)	0.01 (-0.06,0.07)	1.00 (1.00,1.00)
Exit, age 1	rel var	0.04 (0.01,0.08)	0.97 (0.93,1.02)	1.01 (0.98,1.05)	1.00 (1.00,1.00)
	corr	0.33 (0.25,0.41)	0.99 (0.97,1.00)	1.00 (0.99,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.04 (0.00,0.07)	-0.01 (-0.03,0.02)	0.03 (-0.02,0.08)	0.38 (0.34,0.42)
	corr	0.18 (0.02,0.25)	-0.01 (-0.03,0.04)	0.04 (-0.03,0.13)	0.36 (0.30,0.41)
Average size	rel var	0.12 (0.07,0.17)	-0.01 (-0.08,0.02)	0.11 (0.03,0.16)	-0.12 (-0.19,-0.06)
	corr	0.16 (0.09,0.21)	-0.13 (-0.37,0.34)	0.13 (0.03,0.20)	-0.13 (-0.21,-0.07)
		2nd half: 1995–2011			
Employment, age 1	rel var	0.15 (0.13,0.17)	0.20 (0.14,0.28)	0.35 (0.29,0.44)	1.00 (1.00,1.00)
	corr	0.51 (0.44,0.58)	0.87 (0.85,0.89)	0.82 (0.75,0.87)	1.00 (1.00,1.00)
Employment growth, age 2+	rel var	0.08 (0.07,0.08)	-0.01 (-0.02,-0.00)	0.07 (0.05,0.08)	1.00 (1.00,1.00)
	corr	0.44 (0.41,0.47)	-0.10 (-0.14,-0.03)	0.26 (0.20,0.32)	1.00 (1.00,1.00)
Exit, age 1	rel var	0.01 (0.00,0.02)	0.99 (0.98,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
	corr	0.08 (0.02,0.12)	0.99 (0.98,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.05 (0.04,0.06)	0.28 (0.19,0.41)	0.33 (0.23,0.45)	0.96 (0.90,1.03)
	corr	0.41 (0.35,0.46)	0.78 (0.78,0.79)	0.75 (0.72,0.77)	0.90 (0.89,0.90)
Average size	rel var	0.42 (0.38,0.46)	-0.02 (-0.04,-0.00)	0.40 (0.37,0.44)	0.55 (0.52,0.59)
	corr	0.78 (0.75,0.80)	-0.19 (-0.22,-0.02)	0.76 (0.68,0.80)	0.79 (0.75,0.82)

Table C.4: Details of historical variance decomposition: Firms, 1979–2010

		(1) Upside	(2) Downside	(1) & (2)	other
		Full sample			
Employment, age 1	rel var	0.28 (0.26,0.29)	0.20 (0.14,0.26)	0.48 (0.41,0.55)	1.00 (1.00,1.00)
	corr	0.54 (0.51,0.58)	0.63 (0.60,0.68)	0.76 (0.71,0.79)	1.00 (1.00,1.00)
Employment growth, age 2+	rel var	0.01 (0.00,0.01)	-0.05 (-0.06,-0.03)	-0.04 (-0.05,-0.03)	1.00 (1.00,1.00)
	corr	0.02 (0.01,0.04)	-0.28 (-0.29,-0.27)	-0.12 (-0.14,-0.09)	1.00 (1.00,1.00)
Exit, age 1	rel var	-0.01 (-0.02,-0.00)	1.01 (1.00,1.02)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
	corr	-0.11 (-0.14,-0.08)	1.00 (0.98,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.06 (0.04,0.08)	0.25 (0.18,0.34)	0.31 (0.22,0.40)	0.90 (0.86,0.95)
	corr	0.38 (0.35,0.41)	0.56 (0.55,0.58)	0.64 (0.62,0.65)	0.89 (0.88,0.91)
Average size	rel var	0.25 (0.21,0.28)	-0.02 (-0.02,-0.01)	0.23 (0.20,0.26)	0.33 (0.30,0.37)
	corr	0.21 (0.20,0.23)	-0.07 (-0.10,-0.05)	0.21 (0.19,0.22)	0.27 (0.25,0.28)
		1st half: 1979–1994			
Employment, age 1	rel var	0.68 (0.63,0.71)	0.09 (0.06,0.12)	0.76 (0.70,0.83)	1.00 (1.00,1.00)
	corr	0.65 (0.63,0.67)	0.16 (0.11,0.25)	0.71 (0.69,0.73)	1.00 (1.00,1.00)
Employment growth, age 2+	rel var	0.02 (0.01,0.03)	-0.05 (-0.07,-0.04)	-0.03 (-0.05,-0.02)	1.00 (1.00,1.00)
	corr	0.06 (0.04,0.07)	-0.34 (-0.34,-0.33)	-0.10 (-0.14,-0.06)	1.00 (1.00,1.00)
Exit, age 1	rel var	0.03 (0.01,0.05)	0.99 (0.96,1.00)	1.01 (1.00,1.02)	1.00 (1.00,1.00)
	corr	0.22 (0.19,0.26)	0.99 (0.97,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.15 (0.12,0.18)	-0.01 (-0.04,0.01)	0.14 (0.08,0.19)	0.56 (0.52,0.59)
	corr	0.45 (0.40,0.54)	-0.02 (-0.05,0.02)	0.18 (0.10,0.28)	0.58 (0.51,0.65)
Average size	rel var	0.35 (0.29,0.39)	-0.00 (-0.01,0.01)	0.34 (0.30,0.38)	0.21 (0.16,0.26)
	corr	0.26 (0.24,0.28)	-0.02 (-0.04,0.07)	0.26 (0.23,0.28)	0.15 (0.12,0.18)
		2nd half: 1995–2011			
Employment, age 1	rel var	0.24 (0.22,0.25)	0.20 (0.14,0.26)	0.43 (0.37,0.50)	1.00 (1.00,1.00)
	corr	0.68 (0.64,0.74)	0.83 (0.82,0.84)	0.82 (0.79,0.85)	1.00 (1.00,1.00)
Employment growth, age 2+	rel var	-0.04 (-0.04,-0.03)	-0.10 (-0.13,-0.07)	-0.13 (-0.16,-0.11)	1.00 (1.00,1.00)
	corr	-0.17 (-0.19,-0.14)	-0.69 (-0.70,-0.66)	-0.41 (-0.44,-0.37)	1.00 (1.00,1.00)
Exit, age 1	rel var	-0.05 (-0.10,-0.02)	1.05 (1.02,1.10)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
	corr	-0.67 (-0.71,-0.65)	1.00 (0.99,1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)
Entry	rel var	0.03 (0.00,0.06)	0.28 (0.20,0.38)	0.31 (0.22,0.42)	0.82 (0.78,0.89)
	corr	0.36 (0.07,0.48)	0.73 (0.72,0.73)	0.71 (0.69,0.73)	0.91 (0.90,0.91)
Average size	rel var	0.23 (0.20,0.26)	-0.09 (-0.17,-0.04)	0.14 (0.06,0.19)	0.49 (0.39,0.55)
	corr	0.23 (0.21,0.25)	-0.31 (-0.33,-0.26)	0.17 (0.07,0.22)	0.41 (0.37,0.43)

C.8 Robustness of historical variance decomposition

In my baseline estimates, I use employment for businesses aged zero or one year with the upside risk estimated for the entering cohort using the frequencies of the 50 and 100 employee size bins only. The upside risk measure is a moving average of the past two years to match the employment data. I consider the following robustness tests

1. Detrended: Baseline specification, but linear trends removed from all variables, not only risk measure.
2. Age 0 risk and employment: For the same cutoffs, I use age 0 employment and risk without averaging.
3. Age one risk and employment: For the same cutoffs, I use age one risk without averaging.
4. GLS risk: Instead of using a local slope estimator using only the 50- and 100-employee size bins, I use the entire right tail above 50 employees to estimate risk using a weighted least squares estimator. The relative weights are to the inverse frequency of the cells.
5. Baseline, risk 20+: I use the 20- and 50-employee size bins to estimate upside risk.
6. Job creation: Instead of using employment data, I construct employment based on job creation as in Figure C.12: I use age 0 cohort job creation, which is positive by construction, plus age one cohort net job creation, which is negative to obtain each cohort's employment.
7. same, risk 20+: This is the same scenario as above, but using the 20- and 50-employee size bins to estimate upside risk.

Figure C.13 compares the five robustness checks to the baseline scenario. It shows that the baseline result for the variance decomposition is largely robust: The alternative estimates are centered around the baseline estimates.

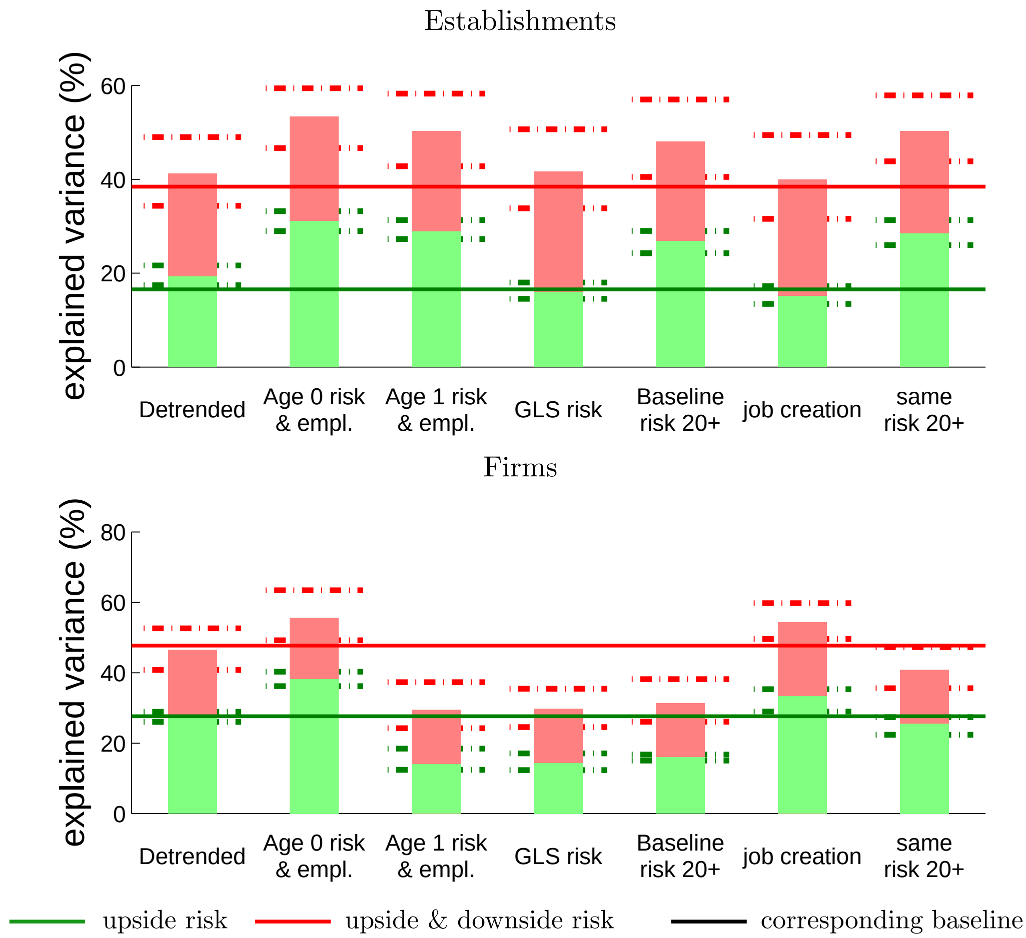


Figure C.13: Fraction of historical variance explained by risk: different specifications

Using different measures for upside risk and employment leaves the main results largely unchanged. Using establishment-level data, about 40% of the historical variance is explained by risk. For firm level data, the results vary between 28% and 68%. Across different specifications, both upside and downside risk contribute to the observed variation in employment.