Under-Diversification and Idiosyncratic Risk Externalities

Felipe S. Iachan, Dejanir Silva, and Chao Zi*

December 24, 2019

Abstract

We study the effects of idiosyncratic uncertainty on asset prices, investment, and welfare. We consider an economy with two main ingredients: i) under-diversification; ii) idiosyncratic risk is endogenous and countercyclical. The equilibrium is subject to underinvestment and excessive aggregate risk-taking. Inefficiencies stem from an idiosyncratic risk externality, as firms do not internalize the effect of their investment decisions on the risk borne by others. Risk externalities depend on an idiosyncratic risk premium and a variance risk premium, and we assess their magnitude empirically. The optimal allocation can be implemented by financial regulation using a tax shield and risk-weighted capital requirements.

Keywords: idiosyncratic uncertainty, under-diversification, risk-taking, pecuniary externalities.

JEL classification: E22, G12, G18.

*Felipe Iachan: FGV EPGE – felipe.iachan@fgv.br; Dejanir Silva: University of Illinois, Urbana-Champaign – dejanir@illinois.edu; Chao Zi: University of Illinois, Urbana-Champaign – chaozi2@illinois.edu.
We would like to thank Antonio Antunes, Victor Duarte, Mahyar Kargar, and seminar participants at the 2019 Lubramacro, UIUC, IDB, Fed Board, SED, and SBE meetings.

1 Introduction

A large body of literature documents under-diversification of idiosyncratic risk.[1] While idiosyncratic risk plays no role in frictionless asset markets, frictions in diversification allow it to affect asset prices, to distort investment and corporate decisions, and to generate economy-wide fluctuations.[2] Moreover, idiosyncratic uncertainty displays a well-documented countercyclicality, so the importance of such effects varies over the business cycle. Nonetheless, little is known about whether and how policymakers could alleviate the inefficiencies created by idiosyncratic risk or respond to its cyclical properties.[3]

In this paper, we study the effects of idiosyncratic uncertainty on asset prices, investment, and welfare. In particular, we consider the (in)efficiency of the equilibrium allocation and its policy implications. We analyze this question in the context of a production asset-pricing model with two main ingredients: (i) under-diversification and (ii) endogenous and countercyclical idiosyncratic risk. In the presence of under-diversification, idiosyncratic risk affects the economy's pricing kernel and, ultimately, investment. Moreover, the quantitative importance of these effects depends on the degree of diversification in the economy. The endogeneity of the countercyclicality of risk plays an important role, as the degree of risk responds to investment decisions and regulation.

Our main result is that the economy is subject to a new form of pecuniary externality, to which we refer as *idiosyncratic risk externalities*: firms do not internalize how their investment decisions affect the level of idiosyncratic risk borne by others. Two implications of these risk externalities are *under-investment* and *excessive aggregate risk-taking* in a laissez-faire economy. Moreover, we derive sufficient statistics for the risk externalities based on asset-price data and quantify the importance of these inefficiencies. Finally, we show how financial regulation can be used to address the inefficiencies caused by idiosyncratic uncertainty.

We consider a two-period model with a unit-mass of firms, investors, and workers. Investment can be allocated across a riskless and a risky technology. The payoff of the risky technology is the only source of aggregate risk and can take two values, either high or low payoff. In the second period, firms combine capital with labor using a Cobb-Douglas production function subject to idiosyncratic productivity shocks. Capital cannot be reallocated once it is installed. Following Gârleanu et al. (2015), investors and firms are located on a circle. Productivity shocks are correlated across firms, with a correlation that decays with the distance between the firms' locations. Average productivity across all locations is non-stochastic, so shocks remain idiosyncratic despite being locally correlated.

Investors choose in the first period how much to consume and an equity portfolio subject to a *limited-participation constraint*. Investors have access only to firms located in a neighborhood of their location. This friction can be interpreted as capturing the fact that investors' portfolios are concentrated

[1]Underdiversification is pervasive for entrepreneurs and outside investors. Himmelberg et al. (2000), e.g. documents that entrepreneurs hold a large fraction of wealth invested in their own companies. Under-diversification in investor's portfolios has been documented by Blume and Friend (1975), Kelly (1995), Polkovnichenko (2005), and Calvet et al. (2007).

[2]See, e.g. Herskovic et al. (2016) for the effects on asset prices, Angeletos and Calvet (2006) and Panousi and Papanikolaou (2012) for the effects on investment, Chen et al. (2010) for the impact on capital structure, and Chen and Strebulaev (2018) for the implications for idiosyncratic risk-taking. Idiosyncratic risk also plays an important role in business cycle research on uncertainty shocks (Bloom 2009), granularity (Gabaix 2011), and networks (Acemoglu et al. 2012). Christiano et al. (2014) identifies uncertainty shocks as the main driver of business cycles.

[3]For instance, the former president of the Dallas Fed, Richard Fisher, highlighted the importance of these issues for policymaking, and the limited attention received until then, in a speech called "Uncertainty Matters" in 2013.

geographically, as documented by Ivković and Weisbenner (2005), or "nearby" firms can be interpreted as those the investor knows about, as in Merton (1987). The important aspect is that investors have access to limited subset of firms. Moreover, the size of the neighborhood, or the length of the arc in the circle, investors have access to acts as a diversification parameter. For example, if an investor has access to the whole circle, she would be able to perfectly diversify the idiosyncratic risk. At the other extreme, if an investor can invest only in a firm at her own location, she would fully bear the idiosyncratic risk of the firm, as, for example, in the entrepreneurship model of Chen et al. (2010). If the investor has access to a positive mass of firms, but less than the full circle, then the investor bears a fraction of the idiosyncratic variance, as the risk is only partially diversified.[4]

Workers play only a role in the second period, when they inelastically supply labor and consume. The significance of having workers in the economy lies in the fact that variations in the cost of labor lead to variations in a firm's operating leverage, inducing *endogenous* movements in idiosyncratic return volatility. The volatility of returns depends on two factors: i) dispersion in the volume produced, determined by the exogenous volatility of productivity, and ii) the profit margin, which is endogenous and varies with economic conditions. For example, in bad times there is weaker demand for labor and lower labor costs, leading to higher profit margins and higher idiosyncratic volatility. Therefore, return risk becomes countercyclical, consistent with the evidence in, for example, Campbell et al. (2001). Moreover, our channel connecting variations in firm-level risk to variations in labor costs is consistent with the recent cross-sectional evidence presented in Donangelo et al. (2019).

This mechanism has important asset-pricing implications. First, the model is able to generate the synchronization of idiosyncratic volatility observed in Herskovic et al. (2016), even without assuming state-dependent productivity dispersion.[5] Second, the model generates a negative premium for exposure to states where idiosyncratic volatility is high, consistent again with the evidence reported in Herskovic et al. (2016). This negative premium results from the stochastic discount factor (SDF) for the economy being the product of a SDF for a representative-agent economy and a term that is increasing with the level of idiosyncratic volatility. This is analogous to the SDF in Constantinides and Duffie (1996), but here the extent of consumption dispersion is related to volatility in firms' returns and the degree of under-diversification in the economy. Given that the SDF increases with idiosyncratic risk, assets that pay off more in states with high volatility command a negative premium. For essentially the same reason, we obtain a positive *idiosyncratic variance risk premium*, that is, a positive difference between expected idiosyncratic variance under the risk-neutral and physical probabilities. A large literature documents a positive aggregate variance risk premium and, in Section 4, we provide evidence of a positive premium for idiosyncratic variance.[6] Moreover, the variance risk premium plays an important role in the analysis of the effects of regulation.

The expected return on the firm can be decomposed in an aggregate risk premium that is pro-

[4]Note the importance of the correlation structure to capture the notion of partial diversification. If productivity were independently distributed across firms then, by the exact law of large numbers (Sun 2006), investors would be able to completely eliminate idiosyncratic risk by investing in any subset of the unit circle with a positive measure.

[5]The model is also consistent with the evidence in Herskovic et al. (2016) that the synchronization of volatility happens both for return volatility and fundamental volatility, measured using the idiosyncratic component in sales.

[6]See, e.g. Bollerslev et al. (2009) and Drechsler and Yaron (2010) for an analysis of the (aggregate) variance risk premium and Zhou (2018) for a recent review of the literature.

portional to the covariance of returns with aggregate consumption, and an *idiosyncratic risk premium* that is proportional to the level of idiosyncratic variance. As in the original model of Merton (1987), we find that idiosyncratic risk commands a positive premium in equilibrium. The price of idiosyncratic risk then depends on the degree of under-diversification. In particular, the price of risk is zero if investors are fully diversified and it is maximized if investors are unable to diversify. Exploring this connection with the idiosyncratic risk premium, we are able to empirically estimate the degree of under-diversification in the economy, a necessary step for our empirical assessment of the welfare implications of idiosyncratic uncertainty.

The asset-pricing implications of uncertainty are transmitted to the real economy through investment decisions that firms make. Idiosyncratic risk leads to a *reduction* in aggregate risk-taking compared with what occurs under perfect markets. This is because investors value the bad state of the world relatively more in the under-diversified economy, given the countercyclicality of volatility and the fact that the SDF is increasing with consumption dispersion. Therefore, idiosyncratic risk leads firms to value the risky technology to a lesser extent, as it is an asset that performs worse in bad times, reducing the amount of aggregate risk-taking. The effect on investment is ambiguous. On the one hand, idiosyncratic risk increases precautionary savings, which translates into an increase in investment. On the other hand, idiosyncratic risk reduces aggregate risk-taking, which reduces precautionary savings. Hence, investment in laissez-faire can be above or below its first-best level.

We next consider the policy implications of the inefficiencies created by idiosyncratic risk. We maintain the assumption that a social planner cannot directly control the degree of diversification of private portfolios. The planner can, however, affect the economy by regulating investment and risk-taking decisions. This constraint reflects the fact that under-diversification may result from limited information or frictions that cannot be directly addressed by the planner.[7]

Our main result is that, in the absence of interventions, the economy is *constrained-inefficient*. In other words, even a planner that is constrained not to directly increase diversification can induce welfare improvements. The inefficiency results from a pecuniary externality in investment decisions. Firms do not internalize the fact that, as they (collectively) increase investment, variable costs rise and operating leverage drops. This effectively reduces the idiosyncratic risk borne by others. A social planner internalizes this additional benefit of investment and perceives underinvestment in the absence of intervention. Similarly, there is excessive aggregate risk-taking, as firms do not internalize how an increase in risk-taking, by shifting resources from bad to good states of the world, increases operating leverage and amplifies idiosyncratic risk when it is especially pronounced. A social planner would then take on less risk than agents in the laissez-faire equilibrium. Note how a planner would like to reduce aggregate risk-taking, despite it being already below the first-best level, and increase investment, regardless of it being above or below the first-best. The direction of the intervention is dictated by the externality, not by a comparison with the first-best. Given that the externality operates through changes in idiosyncratic risk, we refer to this effect as *idiosyncratic risk externalities*.

We consider the effects of small interventions around the laissez-faire equilibrium and show how

[7] Van Nieuwerburgh and Veldkamp (2010) shows that under-diversification may result from an information acquisition problem. Admati et al. (1994) and DeMarzo and Urošević (2006) study how the costs of under-diversification should be balanced against the benefits of better monitoring.

the degree of inefficiency in the economy, or equivalently the gains resulting from regulating investment decisions, can be estimated using asset-price data. In particular, we provide a sufficient statistic for the magnitude of the risk externality in terms of two risk premia. Consider first the impact of increasing investment. We show that the welfare gains depend on the product of the price of idiosyncratic risk and the risk-neutral expectation of the idiosyncratic variance. This quantity can be estimated by combining the idiosyncratic risk premium and the idiosyncratic variance risk premium. Similarly, consider the impact of reducing aggregate risk-taking. We show that the gains of reducing risk-taking depend on the idiosyncratic variance risk premium and the risk-neutral probabilities.[8]

We implement these formulas empirically and show that there are significant welfare gains from correcting risk externalities. We document that investors do not internalize a welfare gain of three cents on each dollar invested. This is equivalent to the social discount rate for the riskless technology being three percentage points lower than the corresponding discount rate for the private sector. We also estimate the gains for reducing aggregate risk-taking. We find that reducing the standard deviation of investment by one unit leads to a welfare gain of 1.2%. This is equivalent to the social planner facing a Sharpe ratio on the risky technology that is four percent smaller than the one for the private sector. In both cases, the magnitude of the idiosyncratic risk externality is significant, suggesting the importance of distortions created by the under-diversification of idiosyncratic risks.

We also consider the design of optimal financial regulation in our environment. We introduce a financial intermediary and show that a tax shield on debt combined with risk-weighted capital requirements are able to increase investment and reduce aggregate risk-taking. This effectively reduces the cost of capital for safe projects and increases the cost of capital for risky ones. Moreover, the magnitude of the optimal tax shield and the optimal risk weights can be related directly to asset prices, analogously to our measurement of the risk externalities.

Literature. Our paper is related to the classical work on under-diversification of Levy (1978), Merton (1987), and Hirshleifer (1988) as well as recent work on the asset-pricing implications of idiosyncratic uncertainty under imperfect risk-sharing, such as Gârleanu et al. (2016), Dou (2016), Di Tella (2017), Silva and Townsend (2019) and Khorrami (2019). Another strand of the literature has focused on the corporate finance implications of idiosyncratic risk, including Chen et al. (2010), Wang et al. (2012), and Chen and Strebulaev (2018). We share with the first strand our focus on how asset prices are affected by idiosyncratic uncertainty and with the second the characterization of how frictions affect investment and risk-taking. In contrast to both lines of work, we emphasize the efficiency properties of the equilibrium and the appropriate regulatory response.[9] In this sense, our approach is similar to that of Di Tella (2019). He considers, however, an environment without endogenous idiosyncratic risk, abstracting from the risk externalities we study here.[10] Our work is also related to the literature on uninsurable income risk, as Constantinides and Duffie (1996), Brav et al. (2002), and Constantinides and Ghosh (2017). Like them, we consider a SDF that depends on countercyclical consumption risk, but we focus instead on the policy implications of endogenous risk in a production economy.

[8]By connecting our sufficient statistics to asset prices, our results show that risk-neutral probabilities are the relevant ones for guiding the design of policy in our environment, consistent with the ideas in Feldman et al. (2015).

[9]In this respect, we are close to Gromb and Vayanos (2002), who also focus on the issue of constrained inefficiency.

[10]Similarly, by abstracting from investment adjustment costs, we ensure that our effects are not caused by the variations of Tobin's Q studied by Di Tella (2019).

2 A model of under-diversification and investment allocation

In this section, we study the implications of under-diversification of idiosyncratic risk for asset pricing and investment decisions. First, we present the environment and then discuss the characterization of the equilibrium. In Section 3, we study the efficiency properties of this economy.

2.1 Environment

We study a finite-horizon economy with two dates, $t = 0, 1$. The economy is populated by workers, investors, and firms, with agents located on a circle of circumference one. Workers play a relevant role only on the last date, when they supply labor and consume. The population of investors consists of a unit-mass of ex-ante identical agents, indexed by $i \in [0, 1)$, who are active in both periods. At date $t = 0$, these agents make consumption and portfolio decisions. There is a unit-mass of ex-ante identical firms indexed by $j \in [0, 1)$. Firms raise equity to finance investment on date zero and pay dividends in period one from the proceeds of the production of final goods.

Uncertainty has both an aggregate and an idiosyncratic component. In particular, at $t = 1$, before production takes place, the aggregate state $s \in \mathcal{S} = \{l, h\}$ is revealed, with $p_s > 0$ representing the probability that each state occurs. We refer to h as the high state, in which production will be endogenously higher, and to state l as the low state. Firm j also learns its idiosyncratic productivity parameter $\theta_j \in \mathbb{R}_+$, which is given by $\theta_j = \Theta e^{-0.5\sigma_\theta^2 + \sigma_\theta \epsilon_j}$, where ϵ_j is normally distributed with a mean of zero and unit variance. The productivity shocks ϵ_j are identically distributed across firms, but are not independent. Their correlation structure is described below. The aggregate state as well as idiosyncratic productivity are public information once realized.

Investment technologies

Firms have access to two investment technologies, $k \in \{0, 1\}$. Technology $k = 0$ delivers $\varphi_s^0 = 1$ units of capital irrespective of the aggregate state, $s \in \mathcal{S}$. We refer to technology $k = 0$ as the *riskless* investment technology.[11] Technology $k = 1$ is a *risky* investment technology and delivers more capital in the good state, that is, its payoff satisfies $\varphi_h^1 > 1 > \varphi_l^1$ and $\mathbb{E}[\varphi_s^1] > 1$. The riskless investment technology corresponds to the standard technology in investment problems without adjustment costs (see e.g. Gomes, 2001), where one unit of investment generates one unit of capital in the following period. The risky technology is subject to capital quality shocks, as in the recent macro-finance literature (see, e.g., Brunnermeier and Sannikov, 2014, and Di Tella, 2017). Importantly, we assume that firms can decide how much to invest in each technology, so the exposure of the economy to aggregate risk is *endogenous* and determined by firms' portfolio choices.

Investment allocation problem

On date $t = 0$, firms must choose how much to invest in each investment technology (I_j^0, I_j^1). The payoff of this investment equals the amount of capital available to the firm in the next period, $K_{s,j} =$

[11] The riskless investment technology is not exposed to risk in the number of effective units of capital it delivers, but the return on each of these units will depend on the firm's idiosyncratic productivity, which is unknown on date $t = 0$.

$\sum_{k=0}^{1} \varphi_s^k I_j^k$. The return on assets (ROA) in period 1 will be given by $R_{s,j}^a \equiv 1 - \delta + \pi_{s,j}$, where δ is the depreciation rate and $\pi_{s,j}$ is the profit per unit of capital generated by the firm, a function of the idiosyncratic productivity θ_j and the aggregate state of the economy s. Let $M_{s,j}$ denote the (average) stochastic discount factor of the firm's shareholders. The problem of the firm can then be written as

$$\max_{I_j^0, I_j^1 \geq 0} \left\{ -\sum_{k=0}^{1} I^k + \mathbb{E}\left[M_{s,j} R_{s,j}^a \sum_{k=0}^{1} \varphi_s^k I_j^k \right] \right\}. \tag{1}$$

The first-order conditions, in an interior solution, imply the investment Euler equations:

$$1 = \mathbb{E}\left[M_{s,j} R_{s,j}^a \varphi_s^k \right],$$

for $k = 0, 1$.

Profit maximization in period 1

Capital cannot be reallocated across firms in period 1, so firm j will operate $K_{s,j}$ units of capital, regardless of its productivity level. This lack of capital reallocation could reflect a financial friction, where the most productive firms are unable to borrow to expand production, or a technological constraint, where capital must be installed in advance, and therefore before the productivity is known. A firm with productivity θ_j and $K_{s,j}$ units of capital hires L workers at wage W_s and produces final goods according to the Cobb-Douglas production function $(\theta_j K_{s,j})^\alpha L^{1-\alpha}$. Each firm chooses how much labor to hire to maximize profits:

$$\max_L \left(\theta_j K_{s,j}\right)^\alpha L^{1-\alpha} - W_s L.$$

The first-order condition for the firm's problem leads to a simple labor demand function,

$$L_{s,j} = \left[\frac{1-\alpha}{W_s} \right]^{\frac{1}{\alpha}} \theta_j K_{s,j}, \tag{2}$$

showing that effective (productivity-adjusted) capital-labor ratios are equalized across firms.

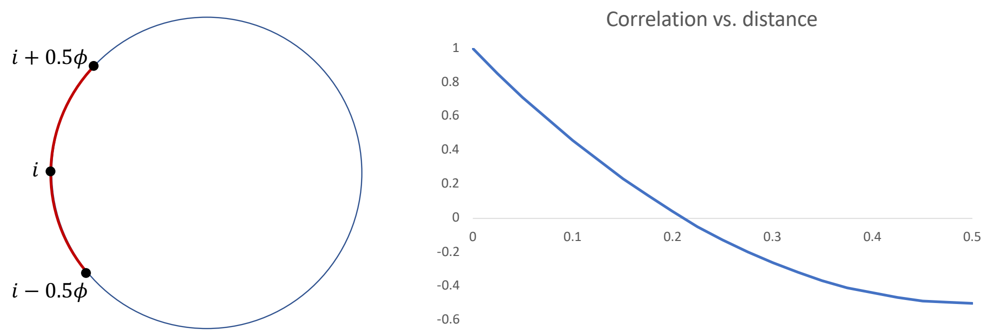
As a consequence of constant returns to scale, the profit function becomes linear in capital and can be written as $\pi_{s,j} K_{s,j}$, where the profit per unit of capital is given by

$$\pi_{s,j} = \alpha \theta_j \left[\frac{1-\alpha}{W_s} \right]^{\frac{1-\alpha}{\alpha}}. \tag{3}$$

Two aspects of expression (3) are worth mentioning. First, profitability is *heterogeneous* across firms. In a frictionless environment, $\pi_{s,j}$ should equal the rental rate of capital for all active firms. As capital does not flow to the most productive firm, firms earn heterogeneous economic rents. Second, the level and dispersion of firms' profitability are *endogenous*. For instance, a reduction in wages reduces variable costs and increases operating leverage, amplifying the effects of changes in productivity.[12]

[12]Formally, variable costs and production are proportional to productivity, $VC_{s,j} = \overline{VC}_s \theta_j$ and $Y_{s,j} = \overline{Y}_s \theta_j$, so the dispersion in profits is $\sigma_{s,\pi} = (\overline{Y}_s - \overline{VC}_s)\sigma_\theta$. Lower wages increase the margin $\overline{Y}_s - \overline{VC}_s$, amplifying the effect of σ_θ.

Figure 1: Participation constraint and correlation structure



The panel on the left describes the participation constraint, where investors can invest only in firms located within distance 0.5ϕ from their location. The panel on the right shows how the correlation varies with the distance.

Correlation structure and the participation constraint

To capture the consequences of under-diversification, we follow Gârleanu et al. (2015) and assume that ϵ_j can be correlated across firms and that the correlation declines with the distance between them. More explicitly, let $d(j, j') = \min\{|j - j'|, 1 - |j - j'|\}$ denote the distance between j and j'. ϵ_j and $\epsilon_{j'}$ are then jointly normal with covariance given by $Cov(\epsilon_j, \epsilon_{j'}) = 1 - 6d(j, j')(1 - d(j, j'))$. Because the maximum distance in the unit-circle is 0.5, the correlation is decreasing with the distance. An important implication of this correlation structure is that productivity shocks are purely idiosyncratic, in the sense that a fully diversified portfolio of these shocks can eliminate the productivity risk, such that h $Var(\int_0^1 \epsilon_j dj) = 0$.[13]

To capture the effects of under-diversification, we assume that investors are subject to *limited participation* in financial markets. Investor i is allowed to invest only in firms located within distance 0.5ϕ of her location, so investors have access to firms on an arc of length ϕ, as indicated in figure 1. The parameter $\phi \in [0, 1]$ controls the degree of under-diversification in this economy. If $\phi = 1$, there is full participation and idiosyncratic risk can be perfectly diversified. If $\phi = 0$, investors are fully invested in a single firm, as in the entrepreneurial models of Chen et al. (2010) and Panousi and Papanikolaou (2012), so they bear all of the idiosyncratic risk. In the case where $0 < \phi < 1$, investors are able to partially diversify the idiosyncratic risk and ϕ measures the degree of diversification.

Formally, a limited-participation constraint takes the following form. Let Ω^i_j denote a cumulative distribution function (cdf) over $j \in [0, 1)$ describing the asset holdings of investor i, that is, the mass of shares of firm j bought by investor i is $d\Omega^i_j$. We do not require Ω^i_j to be continuous.[14]

Let $\mathcal{P}^i = \{j : d(i, j) \leq 0.5\phi\}$ denote the participation set for investor i. The limited participation constraint for investor i is then

$$\int_{\mathcal{P}^i} d\Omega^i_j = 1. \tag{4}$$

[13] Formally, let Z_j denote a Brownian motion in the interval $[0, 1]$ and $B_j = Z_j - jZ_1$ denote a Brownian bridge satisfying $B_0 = B_1 = 0$. The productivity shock is defined as $\epsilon_j \equiv \sqrt{12}\left(B_j - \int_0^1 B_k dk\right)$. We show in the appendix that ϵ_j has a mean of zero, unit variance, the covariance structure described in the text, and that the variance of $\int_0^1 \epsilon_j dj$ is zero.

[14] We assume only that Ω^i_j satisfies the standard properties: it is non-negative, right-continuous, and $\int_{0-}^1 d\Omega^i_j = 1$.

Investor's problem

On date $t = 0$, investors have an endowment of E_0 units of the consumption good and choose how much to consume and how many shares of the various firms to buy, subject to their limited-participation constraint. The investor's problem is

$$\max_{C_0^i, [\Omega_j^i]_{j\in[0,1)}} u\left(C_0^i\right) + \beta \mathbb{E}\left[u\left(C_s^i\right)\right], \tag{5}$$

subject to a non-negativity condition on consumption, the participation constraint (4), and

$$C_s^i = R_s^i(E_0 - C_0^i), \qquad R_s^i \equiv \int_{0-}^{1} \frac{R_{s,j}^a K_{s,j}}{P_j} d\Omega_j^i,$$

where P_j is the price of a share in firm j.

The optimality conditions for this problem are

$$1 = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0^i)} R_s^i\right], \tag{6}$$

and

$$P_j = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0^i)} R_{s,j}^a K_{s,j}\right], \tag{7}$$

for all $j \in \mathcal{P}^i$.

We assume standard isoelastic preferences

$$u(C) = \begin{cases} \frac{C^{1-\gamma}}{1-\gamma} & , if\ \gamma \in \mathbb{R}_+ \setminus \{1\} \\ \log C & , if\ \gamma = 1 \end{cases},$$

where γ represents the constant coefficient of relative risk aversion.

Workers and equilibrium definition

Workers inelastically supply one unit of labor on date 1 and consume their income, i.e.

$$C_s^w = W_s.$$

Let's now define the equilibrium. An *allocation* is given by consumption and portfolio decisions for investors, $\left(C_0^i, [\Omega_j^i]_{j\in[0,1)}\right)$ for $i \in [0,1)$, investment and labor demand decisions for firms, $\left(I_j^0, I_j^1, L_{l,j}, L_{h,j}\right)$ for $j \in [0,1)$, and workers' consumption, (C_l^w, C_h^w). A *competitive equilibrium* is defined as an allocation, asset prices P_j for each firm j, and wages W_s for each state s such that:

i. Consumption and portfolio decisions, $\left(C_0^i, [\Omega_j^i]_{j\in[0,1)}\right)$, solve problem (5) for each $i \in [0,1)$.

ii. Investment decisions solve problem (1) given $M_{s,j} = \frac{1}{\phi}\int_{\{i:j\in\mathcal{P}^i\}} \beta \frac{u'(C_s^i)}{u'(C_0^i)} di$, and labor demand is given by (2).[15]

iii. Worker consumption in each state $s \in \mathcal{S}$ is given by $C_s^w = W_s$.

iv. The asset market clears. Let $S_j \equiv \frac{1}{P_j}\frac{d}{dj}\int_0^1 (E_0 - C_0^i)\Omega_j^i di$ denote the demand for the shares of firm j. Then, for each $j \in [0,1)$,

$$S_j = 1.$$

v. The labor market clears at each $s \in \mathcal{S}$, that is,

$$\int_0^1 L_{s,j} dj = 1.$$

vi. Consumption goods markets clear, that is,

$$\int_0^1 C_0^i di + \sum_{k=0}^{1} I^k = E_0,$$

where $I^k \equiv \int_0^1 I_j^k dj$ for $k = \{0,1\}$, and at each $s \in \mathcal{S}$

$$C^w + \int_0^1 C_s^i di = \int_0^1 (\theta_j K_{s,j})^\alpha L_{s,j}^{1-\alpha} dj + (1-\delta) K_s,$$

where $K_{s,j} = \sum_{k=0}^{1} \varphi_s^k I_j^k$ and $K_s = \int_0^1 K_{s,j} dj$.

2.2 Equilibrium characterization

We consider next the characterization of the equilibrium. We focus on a symmetric equilibrium, where $C_0^i = C_0$, $I_j^k = I^k$, and $P_j = P$. An exact closed-form solution is not available even in the symmetric equilibrium case, but we are able to obtain asymptotic expressions for the case with small idiosyncratic risk.[16] In particular, we consider a first-order perturbation of the equilibrium objects around $\sigma_\theta^2 = 0$. The use of perturbation methods is crucial for managing in a tractable way in our CRRA environment the correlation structure of Gârleanu et al. (2015), originally applied in the context of a model with CARA utility. In particular, we obtain expressions that are analogous to those given by Ito's lemma in continuous time.[17] For instance, we show in appendix A.3 that, for any twice-differentiable function $F(\cdot)$, we obtain

$$\mathbb{E}[F(\sigma_\theta \epsilon_j) - F(0)] = \frac{1}{2}F''(0)\sigma_\theta^2 + o(\sigma_\theta^2), \qquad Var\left[F(\sigma_\theta \epsilon_j)\right] = F'(0)^2 \sigma_\theta^2 + o(\sigma_\theta^2).$$

and all the higher-order central moments are of order $o(\sigma_\theta^2)$.

[15] For ease of exposition, we assume here that $M_{s,j}$ is a simple average of the shareholders' stochastic discount factor. We show in the appendix that our results hold under more general ways of aggregating the investor's SDFs.

[16] See Judd and Guu (2001), for a discussion of asymptotic methods applied to an incomplete markets model.

[17] For a recent application of an under-diversification friction in a continuous-time CRRA model, see Khorrami (2019).

Moreover, the covariance between functions of shocks satisfies

$$Cov(F(\sigma_\theta \epsilon_i), F(\sigma_\theta \epsilon_j)) = F'(0)^2 \sigma_\theta^2 Cov(\epsilon_i, \epsilon_j) + o(\sigma_\theta^2).$$

Applying these Ito-like formulas, we find, for example, that idiosyncratic risk vanishes even for aggregates of non-linear functions of the shocks ϵ_j. In particular, we show that $Var[\int_0^1 \theta_j dj] = 0$, then

$$\int_0^1 \theta_j dj = \Theta,$$

almost surely.

Aggregate production, wages, and returns

Taking the ratio of labor demand (2) for a firm with productivity θ_j and the average labor demand, we obtain $L_{s,j} = \frac{\theta_j}{\Theta}$. We can then compute aggregate output as

$$Y_s \equiv \int_0^1 (\theta_j K_s)^\alpha L_{s,j}^{1-\alpha} dj = (\Theta K_s)^\alpha.$$

Given the Cobb-Douglas production function, the wage is proportional to output

$$W_s = (1-\alpha)\,(\Theta K_s)^\alpha.$$

Plugging the wage into equation (3), we obtain the return on assets

$$R_{s,j}^a = 1 - \delta + \alpha \theta_j \,(\Theta K_s)^{\alpha-1}, \tag{8}$$

which varies with θ and it is decreasing with Θ and K_s.

In equilibrium, the stock price satisfies $P = \sum_{k=0}^1 I^k$, that is, it equals the replacement cost of capital. Therefore the ratio between those two quantities, Tobin's Q, is one.[18] The return on investing in firm j is given by the product of the return on assets and the return on investment

$$R_{s,j} = R_{s,j}^a \frac{\sum_{k=0}^1 \varphi_s^k I^k}{\sum_{k=0}^1 I^k}.$$

Portfolio choice

We now consider the investor's portfolio choice. In a symmetric equilibrium, the expected return is the same across all investors. Intuitively, the investor can then choose her portfolio to eliminate idiosyncratic risk to the extent possible, given the limited-participation constraint. Proposition 1 shows that, in the small-risks case, investors indeed find it optimal to minimize variance.[19]

[18]The fact that Q is equal to one allows us to distinguish the inefficiencies we find from those based on interaction of the price of capital with financial constraints, as in, e.g., He and Kondor (2016) and Jeanne and Korinek (2019).

[19]The proofs for all propositions are provided in the appendix.

Proposition 1 (Portfolio choice.). *Let $\Omega^i_j = \Omega^{i,*}_j + \mathcal{O}(\sigma^2_\theta)$, where $\Omega^{i,*}_j$ denotes the limit of Ω^i_j as σ^2_θ goes to zero. Then, $\Omega^{i,*}_j$ minimizes $Var\left[\int_{0-}^1 \epsilon_j d\Omega^i_j\right]$ subject to the participation constraint (4). The minimal variance is given by*

$$Var\left[\int_{0-}^1 \epsilon_j d\Omega^{i,*}_j\right] = (1-\phi)^3.$$

Moreover, the covariance of ϵ_j, for $j \in \mathcal{P}^i$, with $\epsilon^{i,*} \equiv \int_{0-}^1 \epsilon_k d\Omega^{i,*}_k$ is given by $Cov(\epsilon^{i,*}, \epsilon_j) = (1-\phi)^3$.

Proposition 1 illustrates how the participation constraint affects the level of idiosyncratic risk the investor bears in equilibrium. If $\phi = 1$, the investor holds a fully diversified portfolio, eliminating all of the idiosyncratic risk. If $\phi = 0$, then there is only one firm in the participation set and the investor holds the entirety of the idiosyncratic risk. For $0 < \phi < 1$, the investor effectively bears only a fraction $(1-\phi)^3$ of the risk. Hence, we refer to $\phi_u \equiv (1-\phi)^3 \in [0,1]$ as the *under-diversification parameter*. Moreover, this parameter determines how the investor's portfolio, and ultimately consumption, co-moves with shocks to any firm in the participation set. For this reason, ϕ_u plays a key role in determining the idiosyncratic risk premium, which we study next.

Idiosyncratic risk premium

The log-consumption of investor i, $c^i_s \equiv \log C^i_s$, is given by

$$c^i_s = r^{a,i}_s + k_s,$$

where $r^{a,i}_s \equiv \log \int_{0-}^1 R^a_{s,j} d\Omega^i_j$ is the average ROA on investor i's portfolio and $k_s \equiv \log K_s$.

Define the (log) stochastic discount factor (SDF) for investor i as

$$m^i_s = \log\beta - \gamma(c^i_s - c_0),$$

where $c_0 \equiv \log C_0$.

Given the SDF, we can compute the (shadow) riskless rate. Up to second-order terms, the interest rate is given by the standard expression,[20]

$$r_f \equiv -\log E\left[e^{m^i_s}\right] = -E\left[m^i_s\right] - \frac{\sigma^2_m}{2},$$

where r_f does not vary with i as the distribution of c^i_s is the same for all investors.

Let $r_{s,j} \equiv \log R_{s,j}$ denote the log return on firm j. From the pricing equation for shares (7), we obtain the expected excess return,

$$E\left[r_{s,j}\right] - r_f + \frac{\sigma^2_r}{2} = \gamma Cov\left(c^i_s, r_{s,j}\right),$$

where σ^2_r is the variance of the log-returns.

We can decompose the consumption risk in terms of aggregate and idiosyncratic components. Let $\bar{r}_s = \mathbb{E}_s[r_{s,j}]$ and $\bar{c}_s = \mathbb{E}_s[c^i_s]$ denote the conditional mean of log-returns in state s and the average

[20] This holds for a riskless financial claim to a single unit of $t = 1$ consumption in zero net supply.

log-consumption in the cross-section, respectively. Then,

$$E\left[r_{s,j}\right] - r_f + \frac{\sigma_r^2}{2} = \underbrace{\gamma Cov\left(\bar{c}_s, \bar{r}_s\right)}_{\text{aggregate risk premium}} + \underbrace{\gamma \phi_u \mathbb{E}\left[\sigma_s^2\right]}_{\text{idiosyncratic risk premium}}, \tag{9}$$

where σ_s^2 is the idiosyncratic variance of log-returns in state s.

The risk premium has two components. The first component, which is related to aggregate risk, reflects the usual compensation for the co-movement between aggregate consumption and returns. Given the under-diversification friction, however, investors are also subject to idiosyncratic return risk. This risk requires compensation, which is captured by the second term above. The premium depends on the magnitude of risk $\mathbb{E}[\sigma_s^2]$ as well as the price of risk $\gamma\phi_u$. The price of risk is a function of risk aversion and the *under-diversification parameter* ϕ_u. When $\phi_u = 0$, investors are fully diversified ($\phi = 1$) and the price of idiosyncratic risk is zero. When $\phi_u = 1$, there is no diversification ($\phi = 0$) and the price of risk is at its maximum. Hence, ϕ_u provides not only a measure of under-diversification of investors' portfolios, but also a measure of the required compensation for holding idiosyncratic risk in equilibrium.

Importantly, while the price of idiosyncratic risk is a function of structural parameters, and hence is not directly affected by economic policy, the magnitude of idiosyncratic risk has both endogenous and exogenous components. Given $r_{s,j}^a \approx -\delta + \theta_j \alpha \left(\Theta K_s\right)^{\alpha-1}$, it follows that

$$\log \sigma_s \approx \underbrace{\log \alpha \sigma_\theta}_{\text{exogenous component}} - \underbrace{(1-\alpha)\log\left(\Theta K_s\right)}_{\text{endogenous component}}. \tag{10}$$

Notice that a procyclical quantity of capital generates countercyclical idiosyncratic return risk. The dependence of idiosyncratic return volatility on aggregate variables is consistent with relevant results found in the empirical literature. Campbell et al. (2001) document that idiosyncratic risk is countercyclical. Bekaert et al. (2012) show that average idiosyncratic volatility is correlated across countries and that more than 50% of its variation is explained by aggregate variables. Herskovic et al. (2016) identify a common component in idiosyncratic volatility across firms.[21] These facts can all be explained by our operating-leverage channel, without having to assume shocks to idiosyncratic variance that are correlated across countries or across firms.[22] Moreover, given the endogenous link between idiosyncratic return volatility and aggregate variables, policy interventions can affect the magnitude of idiosyncratic return risk in the economy.

Aggregate risk-taking and investment

We now characterize the magnitude of aggregate risk-taking in the economy, captured by the share invested in the risky technology, $\chi \equiv \frac{I^1}{I^0+I^1}$, and the total level of investment, denoted by $I \equiv I^0 + I^1$. We focus on how idiosyncratic risk affects the overall level and composition of investment. Formally,

[21] Herskovic et al. (2016) document that a similar pattern holds for the idiosyncratic volatility of sales growth, consistent with our result that aggregate variables affect the idiosyncratic volatility of firms' cash flows.

[22] For the relationship between return risk and operating leverage, see Lev (1974). For evidence on this channel, see Novy-Marx (2010) and Donangelo et al. (2019).

we write χ and I as

$$\chi = \chi^* + \hat{\chi}\sigma_\theta^2 + o(\sigma_\theta^2), \qquad I = I^* + \hat{I}\sigma_\theta^2 + o(\sigma_\theta^2).$$

The terms (χ^*, I^*) represent the amount of aggregate risk-taking and investment in an economy without idiosyncratic risk, or alternatively with $\phi_u = 0$. The terms $(\hat{\chi}, \hat{I})$ capture how idiosyncratic risk affects these variables in an economy subject to a diversification friction.

The next proposition describes the sign of the response of idiosyncratic risk on risk-taking and investment. The appendix provides closed-form expressions for both $\hat{\chi}$ and $\hat{I}$.

Proposition 2 (Aggregate Risk-Taking and Investment)**.** *Suppose $\gamma > 1$. Then, $\hat{\chi} < 0$ and the sign of $\hat{I}$ is ambiguous. If firms are constrained to keep $\hat{\chi} = 0$, then $\hat{I} > 0$.*

The result that $\hat{\chi} < 0$ implies that there is *less* risk-taking in the economy that is subject to idiosyncratic risk than in an economy without such risks or with perfect markets. To understand the intuition behind this result, consider the Euler equation

$$0 = \mathbb{E}\left[\mathbb{E}_s\left[(C_s^i)^{-\gamma} R_{s,j}^a\right] \varphi_s^e\right],$$

where $\varphi_s^e \equiv \varphi_s^1 - \varphi_s^0$.

The conditional expectation above can be written as

$$\mathbb{E}_s\left[(C_s^i)^{-\gamma} R_{s,j}^a\right] \approx \overline{C}_s^{-\gamma} \overline{R}_s^a \times \exp\left(\frac{\gamma(\gamma-1)}{2}\phi_u \sigma_s^2\right). \tag{11}$$

The term above acts as a pricing kernel for the investment payoff φ_s^e and it has two components. The first component, $\overline{C}_s^{-\gamma}\overline{R}_s^a$, represents the pricing kernel that would prevail in an economy with complete markets. The second component captures the effects of a precautionary savings motive and, for $\gamma > 1$, it is increasing with the amount of idiosyncratic risk investors effectively bear, $\phi_u \sigma_s^2$. This structure of the pricing kernel is analogous to the one found in Constantinides and Duffie (1996), whose SDF also consists of a representative-agent term and a term that increases with the (state-dependent) consumption dispersion. As in their work, here the countercyclicality of consumption risk plays an important role.[23]

Because the idiosyncratic return risk is countercyclical, $\sigma_l^2 > \sigma_h^2$, the pricing kernel is particularly high in bad times in the case $\phi_u > 0$, so investors dislike risky assets even more in an under-diversified economy. Therefore, idiosyncratic risk reduces aggregate risk-taking under imperfect risk sharing.

The effect on investment $\hat{I}$ is ambiguous, as there are two forces at play. Suppose first that investors cannot adjust the extent of risk-taking. In this case, investment actually increases compared with what occurs in a complete markets economy, as idiosyncratic risk increases precautionary savings. The fact that $\hat{\chi} < 0$ implies, however, that the magnitude of aggregate risk is reduced, pushing savings and investment in the opposite direction. Even though investment may be above or below the first-best benchmark, we show in the next section that there are clear predictions about how a planner should intervene in this economy.

[23] An important distinction between our study and Constantinides and Duffie (1996) is that the consumption countercyclicality is endogenous in our setup, so it potentially responds to policy interventions.

3 Idiosyncratic Risk Externalities

An important aspect of the laissez-faire equilibrium is that the distribution of risk faced by an investor is endogenous, being influenced by both the level and composition of investment. Firms, however, do not internalize how their investment decisions collectively affect the risk born by others. In this section, we illustrate the nature of such effects, which we call *idiosyncratic risk externalities*. Furthermore, we provide sufficient statistics for the welfare gains achieved by regulating investment and aggregate risk-taking. The sufficient statistics are based on two risk premia, an idiosyncratic risk premium and a variance risk premium, which connects the magnitude of risk-sharing inefficiencies in the economy to observable quantities.

3.1 Assessing constrained efficiency

We focus now on the question whether the economy is *constrained-efficient*, that is, whether there are no possible welfare-improving interventions, given the constraints in the economic environment.[24] The economy is obviously inefficient, as risk is not optimally shared across agents, so a planner who could eliminate the under-diversification friction would generate welfare gains. It is much less clear, however, whether interventions that respect the underlying frictions are able to improve welfare. For example, could a planner improve welfare by simply altering the investment decisions, even in the presence of the same degree of under-diversification? In this section, we show indeed that such welfare gains are possible and we provide a characterization of the welfare-improving interventions.

We consider two forms of intervention. The first form increases the overall investment level, while the second form reduces the share invested in the risky technology. We assume that the level and composition of investment can be directly controlled by a social planner and defer the discussion of the implementation of these investment outcomes through financial regulation to Section 5.1.

We characterize a set of Pareto-improving interventions in investment, focusing on their efficiency gains, without the need to assume an explicit social welfare function. To obtain a Pareto improvement, we introduce a fiscal instrument that allows us to keep the utility of workers constant while we search for welfare gains for investors.[25] This instrument consists of a per-unit subsidy on capital, analogous to a depreciation allowance, that is financed by a tax on workers. In the absence of such an instrument, an intervention that, for instance, increases the average capital stock would raise wages and reduce profits, benefiting workers and making investors worse off. Instead, we are interested in the question whether there are net gains after the winners of the intervention compensate its losers, isolating the efficiency gains and avoiding the need to specify preferences for redistribution.

Let Δ parametrize the magnitude of the intervention and let $\tau_s^k(\Delta)$ be the subsidy on capital that is required to maintain workers at their initial consumption level in state s. A general perturbation of

[24] Constrained efficiency has been the standard way to assess the welfare property of economies with frictions since the original work on the efficiency of incomplete markets economies appeared; see Hart (1975), Stiglitz (1982), and Geanakoplos and Polemarchakis (1986).

[25] As the optimal policy analysis from Section 5.2 shows, the constrained inefficiency results do not hinge on a particular sharing of the surplus from the intervention between workers and investors.

investment takes the form

$$I^0(\Delta) = I^0 + \kappa_0\Delta, \qquad I^1(\Delta) = I^1 + \kappa_1\Delta,$$

for some pair of parameters (κ_0, κ_1), and implies a capital at date $t = 1$ given by

$$K_s(\Delta) = K_s + (\kappa_0 + \kappa_1\varphi_s^1)\Delta.$$

Notice that we are able to control the expected value and the riskiness of K_s by adjusting κ_0 and κ_1. The tax that keeps worker's consumption at the laissez-faire level solves

$$C_s^w = (1-\alpha)(\Theta K_s(\Delta))^\alpha - \tau_s^k(\Delta)K_s(\Delta) \implies \tau_s^k(\Delta) = \frac{(1-\alpha)(\Theta K_s(\Delta))^\alpha - C_s^w}{K_s(\Delta)},$$

where C_s^w denotes their consumption in laissez-faire. The ROA for firm j *before* the subsidy is

$$R_{s,j}^a(\Delta) = 1 - \delta + \alpha\theta_j(\Theta K_s(\Delta))^{\alpha-1}.$$

Finally, denote investor i's welfare as a function of the intervention as follows,

$$V(\Delta) = \max_{\Omega_j^i}\left\{u\left(E_0 - \sum_{k=0}^{1} I^k(\Delta)\right) + \beta\mathbb{E}\left[u\left(\left(\int_{0-}^{1} R_{s,j}^a(\Delta)d\Omega_j^i + \tau_s(\Delta)\right)K_s(\Delta)\right)\right]\right\},$$

subject to the limited-participation constraint (4).

If the economy is (constrained) efficient, then $V'(0) = 0$ for any (κ_0, κ_1), so it is not possible to improve welfare by regulating aggregate investment. In contrast, if $V'(0) \neq 0$ for some pair (κ_0, κ_1), then it is possible to design small interventions that generate a welfare gain.

3.2 Underinvestment

For our first main result, we consider a perturbation that increases the expected value of K_s by Δ, while keeping the variance of K_s constant, that is, we set $\kappa_0 = 1$ and $\kappa_1 = 0$.

Proposition 3. *Suppose $\kappa_0 = 1$ and $\kappa_1 = 0$. The marginal gain of increasing Δ, in terms of initial consumption, is given by*[26]

$$\frac{V'(0)}{u'(C_0)} = -(1-\alpha)\mathbb{E}\left[Cov_s\left(\beta\frac{u'(C_s^i)}{u'(C_0)}, R_s^{a,i}\right)\right]$$
$$\approx (1-\alpha)\left[\underbrace{\gamma\phi_u\mathbb{E}\left[\sigma_s^2\right]}_{id.\ risk\ premium} + \gamma\phi_u\underbrace{\left(\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2\right] - \mathbb{E}\left[\sigma_s^2\right]\right)}_{id.\ variance\ risk\ premium}\right] > 0, \tag{12}$$

up to first order in σ_θ^2.

[26] The risk-neutral probabilities satisfy $\mathbb{E}^{\mathbb{Q}}[X_s] = \mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}X_s\right]$ for all random variables X_s, using $\mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}\right] = 1$.

Moreover, for $\gamma \geq 1$, the idiosyncratic variance risk premium is positive, i.e. $\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2\right] - \mathbb{E}\left[\sigma_s^2\right] > 0$.

Proposition 3 shows that there is *underinvestment* in the unregulated economy, that is, the gains obtained by increasing investment are positive. The intuition for this result is the following. An increase in capital stock intensifies competition for labor in the economy, reducing the average profitability of firms. Moreover, this increase in costs affects especially the most productive firms, which are larger and demand more labor. Hence, an increase of the capital stock reduces the dispersion of firms' profitability *ex-post* and the amount of return risk *ex-ante*, as can be seen in equation (10). Firms, however, take prices as given when making their investment decisions, so they do not account for the impact of their actions on the others' risk, generating a (pecuniary) externality. Because the externality operates through changes in the magnitude of idiosyncratic risk, we refer to these effects as *idiosyncratic risk externalities*.

As seen in Section 2.2, the laissez-faire level of investment may be above or below the first-best allocation. Despite this fact, it is always optimal to raise investment in the *second-best* compared with what occurs in the laissez-faire economy. This is because firms do not internalize a potential benefit of investment, the external effect on the risk of others, so there is underinvestment in the economy from the perspective of a social planner. Hence, it is possible that the laissez-faire level of investment is above the first-best level and it remains the case that a further increase in investment achieves a welfare gain.[27]

The magnitude of the inefficiency depends on two distinct risk premia. First, it depends on the magnitude of the *idiosyncratic risk premium*. Given that the idiosyncratic risk premium measures the required compensation an investor demands for taking on idiosyncratic risk, it is intuitive that the magnitude of the welfare gains from reducing such risks, here achieved indirectly through the intervention, is related to the magnitude of this premium. However, one important distinction is that, while we use physical probabilities to compute the expected excess return, the $\mathbb{Q}$-measure is the relevant one with which to compute expected welfare gains. By definition, one dollar in a high-probability state under the $\mathbb{Q}$-measure has a larger impact on welfare than one dollar in a low-probability state. Therefore, risk-neutral probabilities exactly encode the necessary information to perform welfare calculations.

The *idiosyncratic variance risk premium* measures the difference between the expected variance under the risk-neutral and physical probabilities. If idiosyncratic risk were constant across states, this distinction would not be necessary, but given the countercyclicality of return risk, important deviations between the physical and the risk-neutral measure of expected variance can occur. In particular, because the idiosyncratic variance is larger in high marginal utility states, expected variance is higher under the $\mathbb{Q}$-measure, implying a positive idiosyncratic variance risk premium.[28] Notice that the variance risk premium is multiplied by the price of idiosyncratic risk, $\gamma\phi_u$, so its impact on welfare also depends on the degree of diversification. Therefore, the magnitude of the risk externality is proportional to the sum of the idiosyncratic risk premium and an idiosyncratic variance risk premium, adjusted by the

[27] The fact that it may be welfare-improving to move further *away* from the first-best in one dimension is typical of second-best applications, as originally pointed out by Lipsey and Lancaster (1956) in their general theory of second-best.

[28] Much of the literature on the variance risk premium relies on exogenous shocks to the volatility-of-volatility process; see e.g. Bollerslev et al. (2009). In contrast, we are able to endogenously generate the variation in return volatility as well as a positive variance risk premium, despite assuming a constant exogenous volatility of firms' productivity.

degree of diversification.

An alternative way to write expression (12) is $(1-\alpha)\gamma\phi_u\mathbb{E}^{\mathbb{Q}}[\sigma_s^2]$, that is, the welfare gain of the intervention is proportional to the product of the price of idiosyncratic risk, $\gamma\phi_u$, and a term that could be called an *idiosyncratic squared VIX*. Under certain conditions, the squared VIX gives the risk-neutral expectation of the variance for the market as a whole.[29] In contrast, the welfare gains of the intervention are related to the risk-neutral expectation of the idiosyncratic component of firm-level variance.

Another important aspect of formula (12) is that it depends on the labor share $1-\alpha$. Moreover, the inefficiency disappears when $\alpha = 1$. A corollary of this formula is that the economy is constrained-efficient when capital is the only factor of production.

Corollary 1 (Constrained efficiency of the exogenous risk economy)**.** *Suppose $\alpha = 1$. Then, the economy is constrained-efficient, i.e. there is no small intervention on investment or risk-taking that generates a net welfare gain.*

We can explain this result by noting that return risk is completely exogenous when $\alpha = 1$, as can be seen from (10). Investment decisions have no impact on the risk borne by others, so the externality is eliminated and the economy becomes constrained-efficient. Moreover, the economy is also (constrained-) efficient if $\phi_u = 0$. Therefore, our constrained-inefficiency result relies on two key ingredients: i) endogenous return risk, and ii) under-diversification. It is precisely the interaction of these two ingredients that opens the door to welfare-improving interventions.[30]

3.3 Excessive aggregate risk-taking

Our second perturbation consists of an intervention that reduces the share invested in the risky technology. In particular, we choose κ_0 and κ_1 such that the (risk-neutral) standard deviation of capital decreases by Δ, while we keep the total investment unchanged.

Proposition 4. *Suppose $\kappa_0 = \frac{1}{\sqrt{Var^{\mathbb{Q}}[\varphi^1]}}$ and $\kappa_1 = -\frac{1}{\sqrt{Var^{\mathbb{Q}}[\varphi^1]}}$. The marginal gain of increasing Δ, in terms of date $t = 0$ consumption, is given by*

$$\frac{V'(0)}{u'(C_0)} \approx (1-\alpha)\,\gamma\phi_u Cov^{\mathbb{Q}}(\sigma_s^2, \varphi_s^e)\kappa_1$$
$$= (1-\alpha)\gamma\phi_u \underbrace{\left(\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2\right] - \mathbb{E}\left[\sigma_s^2\right]\right)}_{\text{id. variance risk premium}} \zeta > 0, \tag{13}$$

up to the first order in σ_θ^2, with $\zeta \equiv \frac{\sqrt{q_h q_l}}{q_l - p_l}$, where q_s denotes the risk-neutral probability of state $s \in \mathcal{S}$.

Proposition 4 shows that there is *excessive risk-taking* in the unregulated economy. The inefficiency is related to the fact that the risky technology performs poorly when idiosyncratic volatility is high,

[29]The conditions such that the squared VIX equals the risk-neutral expectation of variance, or equivalently the fair strike on a variance swap, likely do not hold in practice, though. See, for example, Martin (2017) for a discussion.

[30]The fact that our results come from this interaction allows us to isolate our channel from previous work on constrained inefficiency in the context of economies with either linear technology, as in Di Tella (2019), or economies without idiosyncratic risk, as in Lorenzoni (2008).

that is, $Cov^{\mathbb{Q}}(\sigma_s^2, \varphi_s^e) < 0$. By shifting resources from bad to good states, risk-taking effectively reduces volatility in good times and increases it in bad times, given the operating-leverage effect. Because bad times are periods in which idiosyncratic risk is already high, aggregate risk-taking imposes a welfare cost on all investors. Hence, private agents take on more aggregate risk than is socially optimal. Note that, even though the risky technology is exposed directly only to aggregate risk, the combination of idiosyncratic risk on profitability and under-diversification leads nevertheless to an inefficient level of risk-taking.

The magnitude of the above effect depends on the price of idiosyncratic risk, $\gamma\phi_u$, and the idiosyncratic variance risk premium. The inefficiency then depends on the *countercyclicality of idiosyncratic risk*, as we would not obtain a positive variance risk premium in the absence of the countercyclicality of risk. We also need the scale factor ζ, which depends on the risk-neutral probabilities, to be able to interpret the intervention as a reduction of one unit in the standard deviation of K_s. Finally, the effect is again proportional to the labor share, as the externality depends on the endogeneity of risk.

3.4 Extensions

In Appendix B.1, we consider three extensions that evaluate the robustness of our main results and offer generalizations. First, we consider an economy with intermediate goods. This extension illustrates how the idiosyncratic risk externality does not exclusively rely on movements in labor costs, but on any variation of marginal costs. The volatility of returns then depends on the relative price of intermediate goods and the externality is present as long as this price moves with the business cycle. If intermediate goods are inelastically supplied, then the expression for the externality is identical to the derived in the baseline model. A positive elasticity of intermediate goods tends to dampen the effect, as part of the adjustment is now coming through quantities instead of prices.

We then consider the case of a constant elasticity of substitution production function. The elasticity of substitution between capital and labor controls how much variations in the capital stock affect firms' marginal costs and ultimately returns. A elasticity of substitution larger than one dampens the effect of the intervention, while low values of the elasticity tends to amplify our effects. The empirical literature typically finds values for the elasticity below one (e.g. Oberfield and Raval, 2014, reports an elasticity of 0.7), suggesting that the gains for the proposed intervention may be actually higher than what the baseline Cobb-Douglas case indicates.

Last, we consider the case in which the participation sets are endogenous, along the lines of Gârleanu et al. (2015). Investors can choose the share of firms ϕ on their participation set subject to paying an increasing and convex utility cost. These costs can be interpreted, for instance, as a cognitive costs related to paying attention to a larger set of firms. We find that our idiosyncratic risk externality is present even in this economy with endogenous participation. The intuition for this result is similar to the one in an envelope theorem. Even though changes in the capital stock may now affect the participation choice, the impact on welfare of these changes in participation is only second-order, given that we start from an optimal participation decision.

4 Measuring Risk Externalities

In this section, we perform the quantification of the risk externalities identified in Propositions 3 and 4. Our goal is to show how asset-pricing data can be used to assess the degree of inefficiency in investment decisions and, therefore, the potential gains of regulation. To perform this exercise, we need to obtain the empirical counterparts of the objects in our expressions for the idiosyncratic risk externalities. To measure the degree of underinvestment, as indicated by the welfare gains represented in expression (12), it is necessary to decompose the idiosyncratic risk premium into the price of idiosyncratic risk, $\gamma\phi_u$, and the magnitude of idiosyncratic risk, $\mathbb{E}[\sigma_s^2]$. We also need a measure of the idiosyncratic variance risk premium and the value of the labor share $1-\alpha$. To measure the degree of excessive risk-taking, as shown in expression (13), it is necessary to identify the risk-neutral probabilities.

We proceed as follows: the price of idiosyncratic risk is obtained by applying standard cross-sectional asset-pricing techniques. In particular, this will allow us to uncover the degree of under-diversification in the economy, ϕ_u, given an estimate of the risk aversion γ. The magnitude of risk is estimated using an EGARCH model for the idiosyncratic variance. The idiosyncratic variance risk premium can be obtained by comparing the corresponding firm-level and market-level premia. Given these quantities, we are able to compute the magnitude of the welfare gains of our proposed interventions.

4.1 Measuring underinvestment

The idiosyncratic risk premium and the variance risk premium

From equation (9), we know that $\gamma\phi_u$ captures the impact of variations in idiosyncratic risk on expected returns, controlling for exposure to aggregate factors. This motivates the following empirical specification:

$$r^e_{i,t+1} = \lambda_0 + \lambda_{id}\mathbb{E}_t[\sigma^2_{i,t+1}] + \boldsymbol{\lambda}'\,\boldsymbol{X}_{i,t} + \epsilon_{i,t+1},\ i = 1, ..., N,\ t = 1, ..., T-1, \tag{14}$$

where $r^e_{i,t+1}$ is the realized excess return on stock i in period $t+1$, $\mathbb{E}_t[\sigma^2_{i,t+1}]$ is the expected variance of the idiosyncratic return in $t+1$ conditional on information in t, and $\boldsymbol{X}_{i,t}$ is a vector of other characteristics that are well-known proxies for a stock's exposure to standard aggregate risk factors. Our primary interest is the slope coefficient λ_{id} for $\mathbb{E}_t[\sigma^2_{i,t+1}]$, which we refer to as the price of idiosyncratic risk. The theory predicts that $\lambda_{id} = \gamma\phi_u$ should be positive, which means that a higher expected idiosyncratic risk is associated with a higher expected excess return. Given a value for γ, we can then use the estimate of λ_{id} to back out the value of the under-diversification parameter ϕ_u.

Addressing the idiosyncratic volatility puzzle. In an influential article, Ang et al. (2006) studied the cross-sectional relationship between idiosyncratic risk and expected returns and found a *negative* price of risk. From the perspective of theory, this result could be a reflection of either not fully controlling for an exposure to aggregate factors or a consequence of mismeasurement in the expected idiosyncratic

Table 1: Summary statistics. This table summarizes monthly stock returns ($r_{i,t}$) and a selection of salient characteristics for a sample of CRSP stocks that are ordinary common shares issued by companies incorporated in the U.S. and listed on the NYSE, the AMEX, or NASDAQ. The sample spans the period running from 1963M07 through 2018M12. The selected characteristics include: $\mathbb{E}_{t-1}[\sigma^2_{i,t}]$, expected idiosyncratic variance estimated by EGARCH models; β^W, Welch (2019) market beta; ME, market capitalization of the issuing firm (converted into real terms using the CPI); BM, book-to-market ratio of the issuing firm; $R_{t-7\to t-2}$, six-month cumulative gross return in the recent past (skip one adjacent month); $TURN$, average monthly turnover; $CVTURN$, coefficient of variation for monthly turnover. Note that some variables are logarithmized following the literature. A 99% winsorization is applied to reduce the influence of outliers.

			Percentiles				
Characteristics	Mean	S.D.	10th	25th	50th	75th	90th
$r_{i,t}$ (%)	1.060	14.010	-13.905	-6.015	0.000	6.977	16.129
$\mathbb{E}_{t-1}[\sigma^2_{i,t}]$ (%)	1.844	3.080	0.253	0.458	0.944	2.032	4.053
β^W	0.801	0.454	0.235	0.455	0.761	1.102	1.418
$ln(ME)$	3.901	2.130	1.194	2.347	3.823	5.384	6.684
$ln(BM)$	-0.493	0.867	-1.569	-0.965	-0.411	0.068	0.493
$R_{t-7\to t-2}$	1.067	0.368	0.680	0.862	1.033	1.213	1.450
$ln(TURN(\%))$	1.649	1.132	0.194	0.853	1.653	2.468	3.118
$ln(CVTURN(\%))$	4.088	0.478	3.475	3.757	4.083	4.395	4.692

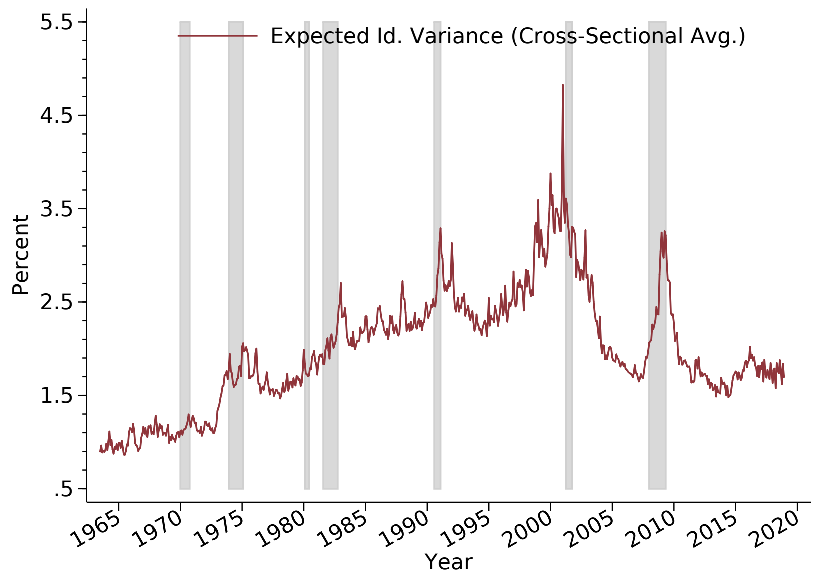
variance. Consider the issue of measuring future expected volatility.[31] Fu (2009) points out the importance of accounting for the fact that volatility is mean-reverting. Assuming mean-reversion, the lagged realized volatility used by Ang et al. (2006) may be an imprecise measure of future expected volatility, biasing the results. Fu (2009) estimates expected future volatility using an EGARCH model and finds a positive price of idiosyncratic risk. Mehra et al. (2019) has recently extended this methodology to allow for time-variation in risk compensation and also finds a positive premium.[32] We follow Fu (2009) and estimate future expected idiosyncratic variance using an EGARCH model.

Sample and variables. Following the convention, we test specification (14) on the cross section of CRSP stocks. Our sample includes stocks that are ordinary common shares issued by companies incorporated in the U.S. and listed on the NYSE, the AMEX, or NASDAQ. We obtain from the CRSP database these stocks' monthly returns as well as other relevant information for the period running from 1963M07 through 2018M12. We measure the expected level of idiosyncratic risk for a stock-month by the conditional variance of the idiosyncratic return; we define the idiosyncratic return as the residual excess return that is unexplained by Fama and French's (1993) three factors. Specifically, we

[31] For studies exploring the other possibility, that idiosyncratic volatility may be correlated with factors omitted in the standard return regressions, see e.g. Boyer et al. (2009), Chen and Petkova (2012), and Duarte et al. (2014).

[32] Similarly, Spiegel and Wang (2007) and Eiling (2013) also use the EGARCH methodology and find a positive price of idiosyncratic risk.

Figure 2: Average expected idiosyncratic variance. This figure displays the month-by-month cross-sectional averages of expected idiosyncratic variance. For each stock, the expected idiosyncratic variance for a month is estimated by an EGARCH model. Shaded areas indicate U.S. recessions identified by NBER.



postulate the following representation of excess returns:

$$r^e_{i,t} = \alpha_i + \beta_{i,mkt} r^e_{m,t} + \beta_{i,smb} SMB_t + \beta_{i,hml} HML_t + \varepsilon_{i,t}, \text{ where } \varepsilon_{i,t} \sim N(0, \hat{\sigma}^2_{i,t})$$
$$\ln \hat{\sigma}^2_{i,t} = a_i + \sum_{j=1}^{p} b_{i,j} \ln \hat{\sigma}^2_{i,t-j} + \sum_{k=1}^{q} c_{i,k} \left\{ \theta \left(\frac{\varepsilon_{i,t-k}}{\hat{\sigma}_{i,t-k}} \right) + \nu \left[\left| \frac{\varepsilon_{i,t-k}}{\hat{\sigma}_{i,t-k}} \right| - \sqrt{\frac{2}{\pi}} \right] \right\}, \tag{15}$$

in which the conditional variance of $\varepsilon_{i,t}$ is our measure of expected idiosyncratic risk; it is represented by an EGARCH model. Following Fu's (2009) procedure, we estimate, for each stock, nine versions of the model with various combinations of p and q as the EGARCH parameters, and we pick the one with the lowest Akaike Information Criterion (AIC). Then for each month we use the selected model to provide a prediction of the idiosyncratic risk conditional on information from the recent past. As shown in Table 1, the median expected idiosyncratic variance in our sample is 0.94% (that is, 9.72% in volatility), similar to that reported in Fu (2009). In Figure 2, we plot, month by month, the cross-sectional averages of expected idiosyncratic variance. One can clearly see evidence of countercyclicality in this series: there are sizeable spikes in almost every recession.

Besides the expected idiosyncratic variance, we also compute a selection of other characteristics for each stock, which include: β^W, the market beta computed via Welch's (2019) approach; ME, the market capitalization of the issuing firm (converted into real terms using the CPI); BM, the book-to-market ratio of the issuing firm; $R_{t-7\to t-2}$, the six-month cumulative gross return in the recent past (skip one adjacent month); $TURN$, the average monthly share turnover; and $CVTURN$, the coefficient of variation for monthly turnover. Detailed definitions of these variables are provided in Appendix C.1.

Table 2: Fama-MacBeth regressions. This table reports the estimation results of Fama-MacBeth regressions specified as $r^e_{i,t+1} = a + \lambda_{ivar}\mathbb{E}_t[\sigma^2_{i,t+1}] + \boldsymbol{\lambda}' \mathbf{X}_{i,t} + \epsilon_{i,t+1}$, where $r^e_{i,t+1}(\equiv r_{i,t+1} - r_{f,t})$ is the return on stock i in excess of the one-month Treasury bill rate in month $t+1$, and $\mathbb{E}_t[\sigma^2_{i,t+1}]$ is the expected idiosyncratic variance in month $t+1$ based on EGARCH prediction. $\mathbf{X}_{i,t}$ is a vector of other characteristics that are known in month t; they include: β^W, Welch (2019) market beta; $ln(ME)$, log market capitalization of the issuing firm; $ln(BM)$, log book-to-market ratio of the issuing firm; $R_{t-6\to t-1}$, past cumulative gross return; $ln(TURN)$, log average monthly turnover; and $ln(CVTURN)$, log coefficient of variation for monthly turnover. In square brackets are Fama and MacBeth (1973) t-statistics with Newey and West (1987) correction (one lag). The sample period is 1963M07 to 2018M12.

	$r^e_{i,t+1}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{E}_t[\sigma^2_{i,t+1}]$				0.35	0.38	0.46	0.45	0.52
				[13.50]	[14.80]	[20.38]	[20.88]	[20.33]
β^W	-0.27	-0.09	0.11		-0.39	-0.57	-0.58	-0.22
	[-1.52]	[-0.44]	[0.71]		[-2.29]	[-3.02]	[-3.28]	[-1.48]
$ln(ME)$		-0.01	-0.13			0.24	0.21	0.09
		[-0.19]	[-3.31]			[6.57]	[6.23]	[2.57]
$ln(BM)$		0.31	0.20			0.49	0.46	0.35
		[6.72]	[4.71]			[11.11]	[10.43]	[8.34]
$R_{t-6\to t-1}$			0.93				0.96	1.01
			[6.32]				[6.72]	[7.04]
$ln(TURN)$			-0.17					-0.31
			[-3.79]					[-7.22]
$ln(CVTURN)$			-0.57					-0.78
			[-11.12]					[-15.73]
a (constant)	0.81	0.87	2.75	-0.07	0.22	-0.36	-1.31	2.51
	[4.87]	[3.52]	[7.52]	[-0.32]	[1.48]	[-1.66]	[-4.77]	[6.91]

Fama-MacBeth regressions. Within this sample of stocks, we estimate (14) via a standard Fama and MacBeth (1973) procedure. Specifically, we perform, month by month, cross-sectional regressions of excess stock returns on expected idiosyncratic variance as well as other characteristics. We then compute time-series averages of the slope coefficients obtained from these cross-sectional regressions, as well as the corresponding Fama-MacBeth t-statistics with Newey and West (1987) correction (one lag). We report these results in Table 2.

We start by replicating some well-documented results in the literature. The results we report in column 1 of Table 2 confirm Fama and French's (1992) finding that market beta alone does not have much explanatory power for average stock returns. In this case the average slope for market beta is negative, contrary to the prediction of standard asset-pricing theory. Column 2 indicates that, when we add size and the book-to-market ratio as explanatory variables, we observe a strong value effect (that is, stocks with high book value of equity relative to their market value tend to bring higher average returns), yet we also observe a weak size effect (that is, big firms tend to have lower stock returns). The slope for market beta remains negative and insignificant. To obtain the results reported in column 3, we further include a measure of past performance as well as two measures of liquidity and its variability.

Now we observe a strong size effect: the slope for $ln(ME)$ is negative and significant. In addition, we also see strong momentum and liquidity effects, as documented by Jegadeesh and Titman (1993) and Chordia et al. (2001), among others. Stocks bringing high returns in the past, displaying low liquidity, or featuring low variability of liquidity tend to bring higher returns. The slope for market beta turns positive but is still insignificant.

Next we turn to the main results. To obtain the results reported in column 4 of Table 2, we use expected idiosyncratic variance alone to explain the cross section of average stock returns. We find that idiosyncratic risk has strong explanatory power for average returns: the slope for $\mathbb{E}_t[\sigma^2_{i,t+1}]$ is positive and 13.50 standard errors away from zero; its magnitude suggests that a one percentage point increase in expected idiosyncratic variance is associated with a 35 basis point increase in average stock return. In the remaining columns, we include other characteristic variables to control for exposure to common risk factors. We find that the explanatory power of idiosyncratic risk becomes even stronger: the slopes for $\mathbb{E}_t[\sigma^2_{i,t+1}]$ are always more than 10 standard errors from zero, and their magnitudes suggest that a one percentage point increase in expected idiosyncratic variance is associated with a 38 to 52 basis point increase in average stock return. In Appendix C.1, we report additional robustness tests. We consider the issue of time variation in the idiosyncratic risk premium. We show that, especially since the 1980s, there is no indication of cyclical variation in the price of risk. This is consistent with our model, in which the price of idiosyncratic risk is equal to the product of γ and ϕ_u, both of which are constant.

In sum, our empirical investigation reveals a strong positive relationship between idiosyncratic risk and average returns. The estimates suggest that a one percentage point increase in expected idiosyncratic variance is associated with a roughly 35-50 basis point increase in average return. Our estimates also suggest that the price of risk for expected idiosyncratic variance is stable and mostly acyclical.

Idiosyncratic variance risk premium. The analysis so far has allowed us to obtain the idiosyncratic risk premium and its decomposition into the price and the magnitude of idiosyncratic risk. To measure the welfare gain in (12), it remains to specify the idiosyncratic variance risk premium. Most of the work on this topic, however, focuses on the variance risk premium for a market index, while it is the variance risk premium associated with the idiosyncratic component that is relevant to our welfare calculation. To isolate this component, we rely on the work of Han and Zhou (2012), who estimated the variance risk premium using stock-level variance (including both aggregate and idiosyncratic components) as well as the market variance risk premium. It turns out that this is all that is necessary to compute the idiosyncratic variance risk premium. To show this, we use the following return decomposition proposed in Campbell et al. (2001), which delivers an additive decomposition of total variance.

Lemma 1 (Variance decomposition)**.** *Let $r_{j,t}$ and $r_{m,t}$ denote the return on firm j and the return on the market, respectively, and define $v_{j,t} \equiv r_{j,t} - r_{m,t}$. Then,*

$$\overline{\sigma}^2_t = \sigma^2_{m,t} + \overline{\sigma}^2_{id,t},$$

where $\overline{\sigma}^2_t$ is the cross-sectional average of individual stock variance, $\sigma^2_{m,t}$ is the market variance, and $\overline{\sigma}^2_{id,t}$ is the cross-sectional average of the variance of the idiosyncratic component $v_{j,t}$.

Table 3: Investment externality parameters.

$\gamma\phi_u$	$\mathbb{E}[\sigma_s^2]$	IRP	VRP	γ	ϕ_u	α
0.35	11.3%	1.8%	3.0%	10	3.5%	0.33

Define the (average) idiosyncratic variance risk premium as $\overline{VRP}_{id,t} \equiv \mathbb{E}_t^{\mathbb{Q}}[\overline{\sigma}^2_{id,t+1}] - \mathbb{E}_t[\overline{\sigma}^2_{id,t+1}]$. Then, from Lemma 1, we can immediately derive $\overline{VRP}_{id,t}$ as

$$\overline{VRP}_{id,t} = \underbrace{\mathbb{E}_t^{\mathbb{Q}}[\overline{\sigma}^2_{t+1}] - \mathbb{E}_t[\overline{\sigma}^2_{t+1}]}_{\overline{VRP}_t} - \underbrace{\left(\mathbb{E}_t^{\mathbb{Q}}[\sigma^2_{m,t+1}] - \mathbb{E}_t[\sigma^2_{m,t+1}]\right)}_{VRP_{m,t}}.$$

Han and Zhou (2012) report that the average value of $\overline{VRP}_t$ over their sample is 5.88%, annualized, while the average of $VRP_{m,t}$ is 2.83%. Therefore, our estimate of the idiosyncratic variance risk premium is $\overline{VRP}_{id} = 5.88\% - 2.83\% \approx 3.05\%$.

The investment externality

Table 3 contains all the elements necessary to compute the idiosyncratic risk externality on riskless investment, $IRE_0 = (1-\alpha)\,[IRP + \gamma\phi_u VRP]$, where IRP is the idiosyncratic risk premium and VRP is the variance risk premium. From our empirical analysis, we found that the price of idiosyncratic risk falls within a range running from 0.35 through 0.5. To be conservative, we choose at lower range of the interval and set $\gamma\phi_u = 0.35$. Even though the price of risk is all we need for this calculation, we can back out ϕ_u using an estimate of the risk aversion γ. Bansal et al. (2016) estimates a value of $\gamma = 9.7$, which we round up to 10. This implies a diversification parameter of $\phi_u = 3.5\%$, meaning that, on average, investors bear only 3.5% of the idiosyncratic variance on stock markets.[33] As our estimate of the idiosyncratic variance, we consider the median idiosyncratic variance reported in Table 1 annualized, that is, $\mathbb{E}[\sigma_s^2] = 11.3\%$. The idiosyncratic variance risk premium is $VRP = 3.0\%$. The value of the idiosyncratic risk externality on investment is then

$$\begin{aligned} IRE_0 &= (1-\alpha)\,[IRP + \gamma\phi_u VRP] \\ &= \frac{2}{3} \times [0.35 \times 11.3\% + 0.35 \times 3.0\%] \approx 3.3\%. \end{aligned} \tag{16}$$

Interpretation. We consider three distinct interpretations of the above number. First, in terms of equivalent wealth gains. An increase in initial wealth of one unit generates a welfare gain of $u'(C_0)$, which, after dividing by the period 0 marginal utility, generates a welfare gain of exactly one unit of consumption. Hence, investors do not internalize a gain worth three cents of wealth for each dollar invested.[34]

The second interpretation of the above number is as a value of insurance. Notice that expression

[33]Using $\phi_u = (1-\phi)^3$, we can back out the value of $\phi \approx 67\%$, which implies that investors have access to roughly two-thirds of the universe of assets.

[34]The value of the capital stock in the US is around \$50 trillion. This means that an increase in capital of one percent, \$560 billion, generates an additional welfare gain of \$16.8 billion.

(12) can be written as

$$IRE_0 = \underbrace{-1+\mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}\right]}_{\text{private trade-off}} - \underbrace{(1-\alpha)\mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}\left(R_s^{a,i}-\overline{R}_s^a\right)\right]}_{\text{externality}}$$
$$= -1+\mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}\left(\alpha R_s^{a,i}+(1-\alpha)\overline{R}_s^a\right)\right].$$

The first term captures the private trade-off which, by the Euler equation, is equal to zero. The planner internalizes an additional effect that acts as insurance: it is negative if the firm's profitability is above average and positive otherwise. The planner effectively perceives the return risk as only a fraction α of what private investors perceive. The externality value of 3% can then be interpreted as a price of three cents for an "insurance policy" of $1-\alpha$ for each dollar of notional value.

The third interpretation is that the social cost of capital is smaller than the private cost. As seen above, the social value of one unit of capital is $Q^{social} = \mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}\left(\alpha R_s^{a,i}+(1-\alpha)\overline{R}_s^a\right)\right] = 1+IRE_0$. A high social value of capital implies an expected return on the investment perceived by the planner that is smaller than the private return. As the expected return on the firm, or equivalently its cost of capital, is related to the amount of capital in the economy, the capital stock seems too low from a planner's perspective. Assuming for simplicity there is no aggregate risk, we can relate the capital stock to the cost of capital using (9):

$$\alpha\Theta^\alpha K^{\alpha-1} \approx \underbrace{r_f+\gamma\phi_u\sigma^2}_{\text{cost of capital }\equiv r^{cc}} +\delta \Rightarrow \frac{\Delta Y}{Y} \approx -\frac{\alpha}{1-\alpha}\frac{\Delta r^{cc}}{r^{cc}+\delta}.$$

The above expression signifies the impact on capital stock of a reduction in the cost of capital. Using the estimate of 18% for the user cost $r^{cc}+\delta$ by Barro and Furman (2018), a reduction of 3% in r^{cc} would imply an increase in aggregate output of 8%.[35] Importantly, this calculation should be interpreted only as indicative of the level of inefficiency at the margin, as our estimate of the externality is local. Nevertheless, the result suggests that frictions related to idiosyncratic risk have important implications for the aggregate economy.

4.2 Measuring excessive risk-taking

We now consider the externality associated with aggregate risk-taking. To quantify expression (13), it remains only to determine $\zeta = \frac{\sqrt{q_h q_l}}{q_l - p_l}$, that is, we need to determine the physical and risk-neutral probabilities of the aggregate states. Given that $\sigma_l^2 > \sigma_h^2$, we associate the low state with periods in which idiosyncratic volatility is above the median and the high state with periods in which the idiosyncratic volatility is below the median.[36] Hence, by definition of the states, we know that $p_l = 0.5$. Using the average idiosyncratic variance, conditional on being above the median value, to estimate σ_l^2,

[35] To put these numbers in perspective, Barro and Furman (2018) expected, as a consequence of the 2017 tax reform, if the provision were made permanent, an expansion of aggregate output of roughly 5%.

[36] Recall that the subscript l denotes the state with the low level of *capital*, which implies, in contrast, a high level of idiosyncratic volatility.

and similarly for σ_h^2, we can back out q_l using the expression for the variance risk premium:

$$VRP = (q_l - p_l)(\sigma_l^2 - \sigma_h^2).$$

From the estimated σ_l^2 and σ_h^2 and the value of the idiosyncratic variance risk premium, we obtain $q_l \approx 0.75$. Given q_l and p_l, we can solve for the remaining parameter $\zeta \approx 1.7$. The value of the idiosyncratic risk externality is then given by

$$\begin{aligned} IRE_1 &= (1-\alpha)\gamma\phi_u \left(\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2\right] - \mathbb{E}\left[\sigma_s^2\right]\right)\zeta \\ &= \frac{2}{3} \times 0.35 \times 3.0\% \times 1.7 \approx 1.2\%. \end{aligned} \tag{17}$$

Interpretation. To understand the intuition behind the above number, notice that we can write the expression for the idiosyncratic risk externality on aggregate risk-taking as follows

$$IRE_1 = -\underbrace{\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_s^i \frac{\varphi_s^e}{\sigma_\varphi}\right]}_{\text{private trade-off}} + \underbrace{(1-\alpha)\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} (R_s^i - \overline{R}_s) \frac{\varphi_s^e}{\sigma_\varphi}\right]}_{\text{externality}},$$

where $\sigma_\varphi \equiv \sqrt{Var[\varphi_s^e]}$.

The term capturing the private trade-off is equal to zero. Expanding the first term to account for covariances, we obtain, after some rearrangements, an expression representing the share invested in the risky technology

$$\chi = \underbrace{\frac{\mathbb{E}[\varphi_s^e]}{\gamma\sigma_\varphi^2}}_{\text{myopic component}} - \underbrace{\left(1 - \frac{1}{\gamma}\right)\left[\frac{Cov(\log \overline{R}_s^a, \varphi_s^e)}{\sigma_\varphi^2} - \frac{\gamma\phi_u}{2}\frac{Cov(\sigma_s^2, \varphi_s^e)}{\sigma_\varphi^2}\right]}_{\text{hedging component}}.$$

Analogous to the financial portfolio decisions studied in Merton (1973), we can divide the share invested in the risky technology into a *myopic* and a *hedging* component. The myopic component captures the usual (static) risk-return trade-off, while the second component captures the fact that the ROA varies across states. Importantly, the covariance between idiosyncratic variance and the payoff of the risky technology is *negative*, consistent with the result expressed in Proposition 2, where we found that the presence of idiosyncratic risk reduces aggregate risk-taking relative to a first-best economy.

In contrast to private agents, a social planner internalizes the fact that an increase in aggregate risk-taking would raise idiosyncratic volatility in bad times and reduce it in good times. This makes $Cov(\sigma_s^2, \varphi_s^2)$ effectively more negative, indicating that the planner would choose a smaller share χ than the one chosen by private agents.

Here the externality can be interpreted as reducing the effective Sharpe ratio perceived by the planner. The planner values the risky investment as if the Sharpe ratio on the risky technology is effectively $\frac{\mathbb{E}[\varphi_s^e]}{\sigma_\varphi} - IRE_\chi$. Given an externality value of 1.2% and Sharpe ratio of, say, 0.30, the social Sharpe ratio is $\frac{0.012}{0.30} = 4\%$ below the private one.

4.3 The dynamics of risk externalities

We have so far considered risk externalities in the context of a two-period model, which has allowed us to derive expressions for the inefficiencies in the simplest possible setting, assessing the importance of these frictions from an unconditional perspective. As the importance of the frictions may vary with the state of the economy, we consider a dynamic extension of our sufficient statistic formulas. Our goal is not to provide the most general dynamic model, but instead a model with the minimal deviation from the environment we have considered so far. For this reason, we consider an overlapping generations version of the two-period model described in Section 2.

Dynamic model. The economy is now populated by a continuum of investors and firms located on the circle of circumference one. Firms are identical to the those described in the baseline model. The payoff of the risky technology φ_s^1 follows a two-state Markov-chain, where the probability of transitioning from state s to state s' is $p_{ss'}$, for $s, s' \in \{l, h\}$. Investors live for two periods, leave no bequests, and start with no wealth.

Our two-period model can then be considered a snapshot of the dynamic economy described above. The endowment of the investor in period 0 is now equal to the labor income $E_s(t) = (1 - \alpha)(\Theta K_s(t))^\alpha$, where $s \in \{l, h\}$ denotes the aggregate state in period t. As in the baseline model, we can consider an intervention that changes investment in period t, but keeps the income of the next generation constant. Hence, the new generation plays the role that workers played in the baseline model. We now focus on the welfare of a generation born when the aggregate state is s

$$V_s(\Delta) = \max_{\Omega_j^i} \left\{ u \left(E_s - \sum_{k=0}^{1} I_s^k(\Delta) \right) + \beta \mathbb{E}_s \left[u \left(\left(\int_{0-}^{1} R_{s',j}^a(\Delta) d\Omega_j^i + \tau_{s'}(\Delta) \right) K_{s'}(\Delta) \right) \right] \right\},$$

where $I_s^k(\Delta)$ denotes the perturbation of investment in technology k analogous to the perturbations discussed in Section 3, and $\tau_{s'}(\Delta)$ denotes the tax required to maintain the same income for the next generation of investors as was the case in the laissez-faire economy.

Proposition 5 (Conditional risk externalities)**.** *Consider the effects of regulating investment decisions in the dynamic economy. Then,*

i. Investment

$$\frac{V_s'(0)}{u'(C_{s,0})} \approx (1-\alpha) \left[\underbrace{\gamma \phi_u \mathbb{E}_s \left[\sigma_{s'}^2 \right]}_{\text{id. risk premium}} + \gamma \phi_u \underbrace{\left(\mathbb{E}_s^{\mathbb{Q}} \left[\sigma_{s'}^2 \right] - \mathbb{E}_s \left[\sigma_{s'}^2 \right] \right)}_{\text{id. variance risk premium}} \right].$$

ii. Aggregate risk-taking

$$\frac{V_s'(0)}{u'(C_{s,0})} \approx (1-\alpha) \gamma \phi_u \underbrace{\left(\mathbb{E}_s^{\mathbb{Q}} \left[\sigma_{s'}^2 \right] - \mathbb{E}_s \left[\sigma_{s'}^2 \right] \right)}_{\text{id. variance risk premium}} \zeta_s > 0,$$

where $\zeta_s \equiv \frac{\sqrt{q_{sh} q_{sl}}}{q_{sl} - p_{sl}}$, *and* $q_{ss'}$ *denotes the risk-neutral probability of state* $s' \in \mathcal{S}$, *conditional on* $s \in \mathcal{S}$.

Figure 3: Risk Externality: Investment

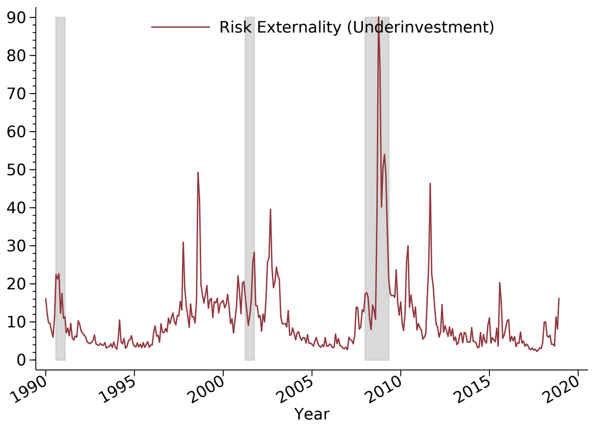
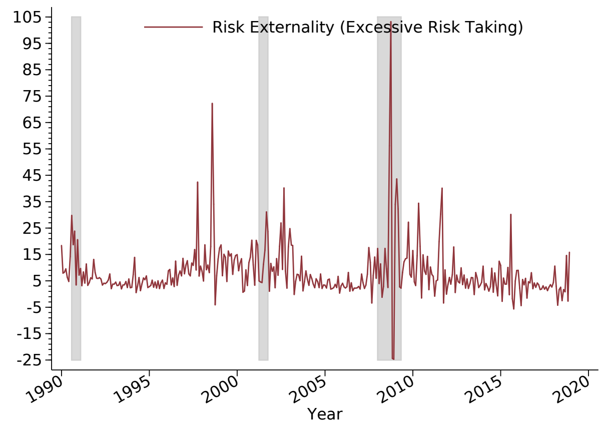


Figure 4: Risk Externality: Risk-Taking



Note: The left panel shows the time series of the conditional risk externality for aggregate investment in basis points. The right panel shows the time series of the conditional externality for aggregate risk-taking in basis points.

The expressions comprising Proposition 5 are conditional versions of our risk-externality formulas. The significance of these expressions is that they allow us to address the question of how fluctuations in idiosyncratic uncertainty affect the efficiency of the economy. The degree of inefficiency fluctuates to the extent that the idiosyncratic risk premium and the variance risk premium vary over time. As can be seen in figure 2, the magnitude of idiosyncratic risk varies substantially over the cycle. Given the stability of the price of idiosyncratic risk, this implies significant variation in the idiosyncratic risk premium over the business cycle. Similarly, the variance risk premium is also time-varying.

Figures 3 and 4 show the time series of the conditional risk externality for investment and aggregate risk-taking. There is substantial variation in the level of the risk externalities, indicating that the inefficiencies are especially more severe in bad times, when idiosyncratic uncertainty is high. In particular, the two externality measures spike during the recent financial crisis, indicating that those are periods in which the discrepancy between the social planner and the private agents in the incentive to invest and take risk is highest.

The time-variation found in the level of externalities suggests the need for countercyclical regulation to address the inefficiencies created by uncertainty risk. An example of such a regulation would be countercyclical capital buffer (CCyB) included in Basel III. We show next that a form of (risk-weighted) capital requirement can be used to address risk externalities.

5 Risk Externalities and Financial Regulation

In this section, we address two questions related to the regulation of risk externalities: *implementation* and *optimal policy*. When deriving the sufficient statistic for the externality discussed in Section 3, we assumed that the planner could directly control investment and risk-taking decisions. In practice, however, these outcomes must be achieved indirectly through regulation. We show that two standard regulatory instruments, a tax shield on debt and risk-weighted capital requirements on financial intermediaries, are capable of implementing the desired allocation. We also consider the optimal level of regulation. We solve the optimal policy problem and show how to relate the optimal level of the regulatory instruments to risk externalities and, ultimately, asset prices.

5.1 Implementation and financial regulation

We assume that the planner controls investment and risk-taking through *financial regulation*. We introduce a continuum of (local) financial intermediaries that raise funds from investors to finance firms. These intermediaries are subject to regulatory constraints, issue debt and equity, and use the proceeds to finance the firm at their own locations. We assume that the intermediary j is in a bilateral relationship with firm j and the terms of the lending contract are determined through bargaining. For the sake of simplicity, we assume that the financial intermediary has all the bargaining power, so firms make no profits in equilibrium, as the rents earned by the firm are extracted entirely by the intermediary. Given that firms make no profits, we simply assume that they are entirely bank-financed and that the investors then choose a portfolio of financial firms that is subject to the limited-participation constraint (4).[37]

Financial intermediaries' problem. Each intermediary $j \in [0,1)$ maximizes the value of equity. It also issues (riskless) deposits to investors, in quantity D_j. Intermediaries receive a subsidy on deposits of τ_d, which can be interpreted as a tax shield. Let P_d denote the price investors pay on the deposit (implying an interest rate of $1/P_d$), so that the intermediary receives $P_d(1+\tau_d)$ for each unit of deposit.

As intermediaries have all the bargaining power, they maximize the surplus of the relationship with the firm. Hence, the intermediary chooses the level of investment to maximize the operational profit generated by the firm, net of the intermediaries' borrowing costs. Formally, the intermediary solves the problem

$$\max_{D_j, I_j^0, I_j^1 \geq 0} \left\{ P_d \left(1+\tau_d\right) D_j - \sum_{k=0}^{1} I_j^k + \mathbb{E}\left[M_{s,j} \left(R_{s,j}^a \sum_{k=0}^{1} \varphi_s^k I_j^k - D_j \right) \right] \right\}, \tag{18}$$

subject to

$$D_j \leq (1-\delta) \sum_k I_j^k \varphi_L^k, \qquad \sum_k I_j^k - P_d D_j \geq \sum_k \omega^k I_j^k. \tag{19}$$

In the objective function, the difference between the first two terms represents the amount of equity raised by the intermediary in the first period. The last term corresponds to the surplus generated by the firms, net of deposits, discounted by the shareholder's SDF. The first constraint in (19) guarantees that deposits are riskless. The second one is a regulatory constraint, a *risk-weighted capital requirement*, according to which equity must exceed risk-weighted assets, given weights ω^k for $k = 0, 1$.

Investment and risk-taking wedges. Consider the capital structure choice of the intermediary. Suppose initially that the regulatory constraint is not binding. Given our assumption that deposits are

[37] The assumption that the intermediary has all the bargaining power simplifies the exposition, but it is not essential for the argument. We could have assumed instead that firms have some bargaining power, so they would make profits. This would require, however, to characterize the capital structure of both intermediaries and non-financial firms. As this additional layer of complexity is not necessary for our implementation result, we abstract from these features.

riskless, there is no cost of default, unlike in the standard trade-off theory of capital structure. As a consequence, the intermediary would choose the maximum amount of debt to obtain the benefits of the tax shield. However, as we introduce the capital requirement, another trade-off emerges: one between the tax benefit and the tightening of the capital requirement constraint. We show in the appendix that, given the tax shield $\tau_d \geq 0$ and the level of debt, we obtain the following distortion in the Euler equation:

$$\mathbb{E}\left[M_{s,j} R^a_{s,j} \frac{K_s}{I - P_d \tau_d D_j}\right] = 1.$$

The tax benefit essentially reduces the cost of investment, creating a wedge in the investment Euler equation. Note that the risk weight does not directly affect the equation above. In contrast, it has a direct impact on the Euler equation for the risky investment,

$$\mathbb{E}\left[M_{s,j} R^a_{s,j} \varphi^e_s\right] = (\omega^1 - \omega^0)\tau_d P_d.$$

Imposing an additional risk weight on risky assets, $\omega^1 > \omega^0$, tends to reduce the intensity of risk-taking in the economy. By reducing risk-taking, intermediaries increase the covariance of φ^e_s with the SDF, as it becomes less negative, until it matches with the right-hand side. The term $P_d \tau_d$ captures the *shadow cost of the regulatory constraint*, as the intermediary optimally balances this shadow cost with the tax benefit.

We show in Appendix D that a planner can use the tax shield, τ_d, and the risk weights, (ω_0, ω_1), to solve the implementation problem. In particular, this result establishes that any allocation that is feasible, constrained in its risk-sharing by limited participation, and features both implicit subsidies to investment and implicit taxes on risk-taking can be implemented as an equilibrium of an economy in which debt is subsidized by a tax shield and a risk-weighted capital requirement constraint is imposed on intermediaries.

5.2 Optimal policy

We turn now to the design of the optimal policy. We seek first to characterize the properties of the (constrained) optimal allocation and then build on the implementation results from the previous section to characterize how a tax shield and a risk-weighted capital requirement can support this allocation in equilibrium. Relative to an unregulated economy, the planner internalizes changes in idiosyncratic risk that would be ignored by private agents, and the magnitude of these external effects are related to the optimal level of the policy instruments.[38]

The key constraint imposed on the planner is limited participation in idiosyncratic risk-sharing. Moreover, we assume that the planner has no instrument with which to distort the portfolio allocation of investors, even among the assets satisfying the limited participation condition. We show, however, this is not a relevant constraint. We consider a relaxed version of the planner's problem, where only the participation constraint is imposed, and then show that it is not optimal to distort portfolio decisions.

[38]The optimal allocation emerging in this section deviates from private optimization in ways that are reminiscent of the perturbation arguments in Section 3.

We write the relaxed planning program as

$$\max_{I,\chi,\Omega_j^i,\{T_s^w\}_s} u\left(E_0 - I\right) + \beta \mathbb{E}\left[u\left(R_j^{a,i} K_s + T_s^w\right)\right], \tag{20}$$

subject to the limited participation constraint (eq. 4) and

$$E\left[u^w\left(\left(1-\alpha\right)\left(\Theta K_s\right)^{\alpha} - T_s^w\right)\right] \geq \underline{u}^w, \tag{21}$$

where $R_{j,s}^a = 1 - \delta + \alpha\theta_j\left(\Theta K_s\right)^{\alpha-1}$, and $K_s = \left(1 + \chi\varphi_s^e\right) I$.

In the above relaxed planning problem, all constraints on feasibility and the distribution of consumption across agents are taken into account. Additionally, Equation (4) imposes the same limited participation in idiosyncratic risk-sharing as before, while Constraint (21) guarantees that workers receive some arbitrary utility level, given by the parameter $\underline{u}^w$. By varying this parameter, along with the lump-sum transfer T_s^w, one can trace out a (constrained) Pareto frontier between workers' and investors' expected utility. The solution to this problem is characterized in the following proposition.

Proposition 6 (Optimal Policy)**.** *The necessary first-order conditions of Problem (20) can be summarized as:*

i. A planner's investment Euler equation:

$$1 = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_{s,j}^a \frac{K_s}{I(1 - IRE_I)}\right], \tag{22}$$

where $IRE_I \approx (1-\alpha)\gamma\phi_u\left[(1-\chi)\mathbb{E}^{\mathbb{Q}}[\sigma_s^2] + \chi\mathbb{E}^{\mathbb{Q}}[\sigma_s^2\varphi_s^1]\right]$.

ii. A planner's risky technology Euler equation:

$$\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_s^{a,i} \varphi_s^e\right] = IRE_\chi, \tag{23}$$

where $IRE_\chi \approx -(1-\alpha)\gamma\phi_u Cov^{\mathbb{Q}}(\sigma_s^2, \varphi_s^e)$.

iii. An optimal portfolio condition, stating that, for all (i, j) *such that* $j \in \mathcal{P}^i$

$$\mathbb{E}\left[u'\left(C_s^i\right) R_{j,s}^a K_s\right] = \mathbb{E}\left[u'\left(C_s^i\right) R_s^{a,i} K_s\right].$$

Proposition 6 provides a characterization of the optimal allocation. The main feature of the solution is that the wedges in the Euler equations for investment and for the share invested in the risky technology depend on terms capturing idiosyncratic risk externalities, analogous to those in Propositions 3 and 4. The intuition for those terms is the same as before: the planner internalizes the impact of investment decisions on the level of idiosyncratic risk. The third condition gives the planner's optimal portfolio condition, which coincides with the condition for private investors (7). Therefore, the optimal policy consists of correcting the investment decisions instead of distorting investors' trading behavior.

An important feature of the solution is that the wedges can be directly related to the two regulatory instruments available to the planner, the tax shield and the risk weights. Comparing the investment

Euler equation for the financial intermediary with the corresponding one for the planner, we obtain

$$P_d d \tau_d = IRE_I, \qquad d \equiv \frac{D}{I}.$$

Hence, the tax benefit, per unit of investment, should equal the risk externality on investment, which is given by the weighted average of the externality on the two technologies. By matching the tax benefit to the risk externality, the planner induces the financial intermediaries to internalize the effects that private agents do not take into account in the laissez-faire equilibrium. Moreover, all the elements required to estimate the tax benefit can be recovered directly from the data, as was illustrated in Section 4. This aspect contrasts with alternative approaches to the analysis of financial regulation, which typically rely more heavily on the calibration and numerical solution of an economic model.

Comparing the Euler equation for the share invested in the risky technology for the financial intermediary and for the planner, we obtain

$$(\omega^1 - \omega^0)\tau_d P_d = IRE_\chi.$$

The above expression connects the risk weights, the tax benefit on debt, and the risk externality on aggregate risk-taking. The term on the left-hand side captures the effect of the regulation on the risk-taking decision. The term $\omega_1 - \omega_0$ corresponds to the extent to which an increase in the share of the risky technology tightens the regulatory constraint, and $\tau_d P_d$ captures the shadow cost of the regulatory constraint. Given that the regulatory cost is an important part of the choice of capital structure, the shadow cost of the regulatory constraint must equalize the tax benefit of debt. The right-hand side captures the externality perceived by the social planner. By matching the effective regulatory cost of the risky technology with the corresponding externality, the planner induces financial intermediaries to take the appropriate degree of risk from a social perspective.

An advantage of our expression for the risk weight is again the fact that it can be estimated directly from the data. In our empirical exercise in Section 4, we connected the risk externality to the idiosyncratic variance risk premium, in the context of our simple two-state model. Our formulas hold more generally, though, and one could apply the same expressions on environments with several assets and aggregate states. In particular, the relative risk weight on two assets can be determined by the expression

$$\frac{\omega^k - \omega^0}{\omega^{k'} - \omega_0} = \frac{Cov^{\mathbb{Q}}(\sigma_s^2, \varphi_s^{e,k})}{Cov^{\mathbb{Q}}(\sigma_s^2, \varphi_s^{e,k'})},$$

for $k, k' = 1, 2, \ldots, K$, where K is the number of risky assets.

The expression above provides a tight connection between data on asset prices and the optimal regulatory risk weight, which can be used to guide financial regulation.

6 Conclusion

In this paper, we study the impact of portfolio diversification frictions on asset prices, investment, and welfare. We consider a production asset-pricing model where investors hold under-diversified portfolios and idiosyncratic return risk is endogenous and countercyclical. We show that, absent intervention, this economy is *constrained inefficient*, featuring underinvestment and excessive aggregate risk-taking. Our main contribution lies in identifying these inefficiencies and connecting their magnitudes to sufficient statistics, whcih can be measured directly in the data. In particular, these statistics are derived from two risk premia: an idiosyncratic risk premium and an idiosyncratic variance risk premium.

We find a significant impact of idiosyncratic uncertainty on welfare and also consider the optimal financial regulation. The optimal allocation can be implemented using two instruments: a tax shield on debt and risk-weighted capital requirements on financial intermediaries. Intuitively, the tax shield on debt stimulates an increase in investment levels, while the appropriate risk weights control risk-taking. The time-varying behavior of these inefficiency measures can provide further guidance to regulators. For instance, given that the measures of inefficiencies are countercyclical, they can be used to inform the implementation of a countercyclical capital buffer.

Our model can be extended in several other directions in future research. For instance, there is extensive work on limited international risk-sharing. Imperfect diversification across international markets may lead to risk externalities and inefficiencies similar to the ones we found in this paper. Additionally, the financial intermediary considered here is not subject to any friction, other than the one imposed by regulation. An interesting research direction is to consider the role of risk externalities in a setting where the balance sheets of intermediaries play an important role, as in the recent intermediary asset-pricing literature (see, e.g., He and Krishnamurthy 2013). Given the importance of financial intermediaries in determining asset prices, this could be another example of how asset-pricing information may be directly relevant to the design of financial regulation.

References

Acemoglu, Daron, Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "The network origins of aggregate fluctuations," *Econometrica*, 2012, *80* (5), 1977–2016.

Admati, Anat R, Paul Pfleiderer, and Josef Zechner, "Large shareholder activism, risk sharing, and financial market equilibrium," *journal of Political Economy*, 1994, *102* (6), 1097–1130.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang, "The cross-section of volatility and expected returns," *The Journal of Finance*, 2006, *61* (1), 259–299.

Angeletos, George-Marios and Laurent-Emmanuel Calvet, "Idiosyncratic production risk, growth and the business cycle," *Journal of Monetary Economics*, 2006, *53* (6), 1095–1115.

Bansal, Ravi, Dana Kiku, and Amir Yaron, "Risks for the long run: Estimation with time aggregation," *Journal of Monetary Economics*, 2016, *82*, 52–69.

Barro, Robert J and Jason Furman, "Macroeconomic effects of the 2017 tax reform," *Brookings papers on economic activity*, 2018, *2018* (1), 257–345.

Bekaert, Geert, Robert J Hodrick, and Xiaoyan Zhang, "Aggregate idiosyncratic volatility," *Journal of Financial and Quantitative Analysis*, 2012, *47* (6), 1155–1185.

Bloom, Nicholas, "The impact of uncertainty shocks," *Econometrica*, 2009, *77* (3), 623–685.

Blume, Marshall E and Irwin Friend, "The asset structure of individual portfolios and some implications for utility functions," *The Journal of Finance*, 1975, *30* (2), 585–603.

Bollerslev, Tim, George Tauchen, and Hao Zhou, "Expected stock returns and variance risk premia," *The Review of Financial Studies*, 2009, *22* (11), 4463–4492.

Boyer, Brian, Todd Mitton, and Keith Vorkink, "Expected idiosyncratic skewness," *The Review of Financial Studies*, 2009, *23* (1), 169–202.

Brav, Alon, George M Constantinides, and Christopher C Geczy, "Asset pricing with heterogeneous consumers and limited participation: Empirical evidence," *Journal of Political Economy*, 2002, *110* (4), 793–824.

Brunnermeier, Markus K and Yuliy Sannikov, "A macroeconomic model with a financial sector," *American Economic Review*, 2014, *104* (2), 379–421.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, "Down or out: Assessing the welfare costs of household investment mistakes," *Journal of Political Economy*, 2007, *115* (5), 707–747.

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.

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

Chen, Zhanhui and Ralitsa Petkova, "Does idiosyncratic volatility proxy for risk exposure?," *The Review of Financial Studies*, 2012, *25* (9), 2745–2787.

Chen, Zhiyao and Ilya A Strebulaev, "Macroeconomic Risk and Idiosyncratic Risk-taking," *The Review of Financial Studies*, 2018, *32* (3), 1148–1187.

Chordia, Tarun, Avanidhar Subrahmanyam, and V.Ravi Anshuman, "Trading activity and expected stock returns," *Journal of Financial Economics*, 2001, *59* (1), 3 – 32.

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

Constantinides, George M and Anisha Ghosh, "Asset pricing with countercyclical household consumption risk," *The Journal of Finance*, 2017, *72* (1), 415–460.

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

DeMarzo, Peter M and Branko Urošević, "Ownership dynamics and asset pricing with a large shareholder," *Journal of political Economy*, 2006, *114* (4), 774–815.

Donangelo, Andres, Francois Gourio, Matthias Kehrig, and Miguel Palacios, "The cross-section of labor leverage and equity returns," *Journal of Financial Economics*, 2019, *132* (2), 497–518.

Dou, Winston Wei, "Embrace or fear uncertainty: Growth options, limited risk sharing, and asset prices," *Unpublished working paper, MIT*, 2016.

Drechsler, Itamar and Amir Yaron, "What's vol got to do with it," *The Review of Financial Studies*, 2010, *24* (1), 1–45.

Duarte, Jefferson, Avraham Kamara, Stephan Siegel, and Celine Sun, "The systematic risk of idiosyncratic volatility," *Available at SSRN 1905731*, 2014.

Eiling, Esther, "Industry-specific human capital, idiosyncratic risk, and the cross-section of expected stock returns," *The Journal of Finance*, 2013, *68* (1), 43–84.

Fama, Eugene F. and James D. MacBeth, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 1973, *81* (3), 607–636.

__ **and Kenneth R. French**, "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 1992, *47* (2), 427–465.

__ **and** __ , "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 1993, *33* (1), 3 – 56.

Feldman, Ron, Ken Heinecke, Narayana Kocherlakota, Sam Schulhofer-Wohl, and Tom Tallarini, "Market-based probabilities: a tool for policymakers," *Federal Reserve Bank of Minneapolis*, 2015.

Fu, Fangjian, "Idiosyncratic risk and the cross-section of expected stock returns," *Journal of Financial Economics*, 2009, *91* (1), 24 – 37.

Gabaix, Xavier, "The granular origins of aggregate fluctuations," *Econometrica*, 2011, *79* (3), 733–772.

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

_ , _ , and _ , "Impediments to financial trade: Theory and applications," Technical Report, National Bureau of Economic Research 2016.

Geanakoplos, John D and Heraklis M Polemarchakis, "Existence, regularity, and constrained suboptimality of competitive allocations," *Essays in Honor of Kenneth J. Arrow: Volume 3, Uncertainty, Information, and Communication*, 1986, *3*, 65.

Gomes, Joao F, "Financing investment," *American Economic Review*, 2001, *91* (5), 1263–1285.

Gromb, Denis and Dimitri Vayanos, "Equilibrium and welfare in markets with financially constrained arbitrageurs," *Journal of financial Economics*, 2002, *66* (2-3), 361–407.

Han, Bing and Yi Zhou, "Variance risk premium and cross-section of stock returns," *unpublished paper, University of Texas at Austin*, 2012.

Hart, Oliver D, "On the optimality of equilibrium when the market structure is incomplete," *Journal of economic theory*, 1975, *11* (3), 418–443.

He, Zhiguo and Arvind Krishnamurthy, "Intermediary asset pricing," *American Economic Review*, 2013, *103* (2), 732–70.

_ and Péter Kondor, "Inefficient investment waves," *Econometrica*, 2016, *84* (2), 735–780.

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

Himmelberg, Charles P, R Glenn Hubbard, and Inessa Love, "Investment, protection, ownership, and the cost of capital," *National Bank of Belgium Working Paper*, 2000, (25).

Hirshleifer, David, "Residual risk, trading costs, and commodity futures risk premia," *The Review of Financial Studies*, 1988, *1* (2), 173–193.

Ivković, Zoran and Scott Weisbenner, "Local does as local is: Information content of the geography of individual investors' common stock investments," *The Journal of Finance*, 2005, *60* (1), 267–306.

Jeanne, Olivier and Anton Korinek, "Managing credit booms and busts: A Pigouvian taxation approach," *Journal of Monetary Economics*, 2019, *107*, 2–17.

Jegadeesh, Narasimhan and Sheridan Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *The Journal of Finance*, 1993, *48* (1), 65–91.

Judd, Kenneth L and Sy-Ming Guu, "Asymptotic methods for asset market equilibrium analysis," *Economic theory*, 2001, *18* (1), 127–157.

Kelly, Morgan, "All their eggs in one basket: Portfolio diversification of US households," *Journal of Economic Behavior & Organization*, 1995, *27* (1), 87–96.

Khorrami, Paymon, "The Risk of Risk-Sharing: Diversification and Boom-Bust Cycles," 2019.

Lev, Baruch, "On the association between operating leverage and risk," *Journal of financial and quantitative analysis*, 1974, *9* (4), 627–641.

Levy, Haim, "Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio," *The American Economic Review*, 1978, *68* (4), 643–658.

Lipsey, Richard G and Kelvin Lancaster, "The general theory of second best," *The review of economic studies*, 1956, *24* (1), 11–32.

Lorenzoni, Guido, "Inefficient credit booms," *The Review of Economic Studies*, 2008, *75* (3), 809–833.

Martin, Ian, "What is the Expected Return on the Market?," *The Quarterly Journal of Economics*, 2017, *132* (1), 367–433.

Mehra, Rajnish, Sunil Wahal, and Daruo Xie, "Is Idiosyncratic Risk Conditionally Priced?," Technical Report, National Bureau of Economic Research, Inc 2019.

Merton, Robert C, "An intertemporal capital asset pricing model," *Econometrica*, 1973, *41* (5), 867–887.

_ , "A simple model of capital market equilibrium with incomplete information," *The journal of finance*, 1987, *42* (3), 483–510.

Newey, Whitney K. and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 1987, *55* (3), 703–708.

Nieuwerburgh, Stijn Van and Laura Veldkamp, "Information acquisition and under-diversification," *The Review of Economic Studies*, 2010, *77* (2), 779–805.

Novy-Marx, Robert, "Operating leverage," *Review of Finance*, 2010, *15* (1), 103–134.

Oberfield, Ezra and Devesh Raval, "Micro data and macro technology," Technical Report, National Bureau of Economic Research 2014.

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

Polkovnichenko, Valery, "Household portfolio diversification: A case for rank-dependent preferences," *The Review of Financial Studies*, 2005, *18* (4), 1467–1502.

Silva, Dejanir and Robert Townsend, "Risk-taking over the Life Cycle: Aggregate and Distributive Implications of Entrepreneurial Risk," 2019.

Spiegel, Matthew I and Xiaotong Wang, "Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk," 2007.

Stiglitz, Joseph E, "The inefficiency of the stock market equilibrium," *The Review of Economic Studies*, 1982, *49* (2), 241–261.

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

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

_ , "Optimal regulation of financial intermediaries," *American Economic Review*, 2019, *109* (1), 271–313.

Wang, Chong, Neng Wang, and Jinqiang Yang, "A unified model of entrepreneurship dynamics," *Journal of Financial Economics*, 2012, *106* (1), 1–23.

Welch, Ivo, "Simpler Better Market Betas," 2019. Working Paper.

Zhou, Hao, "Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty," *Annual Review of Financial Economics*, 2018, *10*, 481–497.

A Appendix for Section 2

A.1 Modeling the correlation structure using Brownian bridges

We now describe in detail the definition of the productivity shocks and the corresponding correlation structure. The results presented in this subsection follow closely the work of Gârleanu et al. (2015) and are presented for the sake of completeness. Let $Z_i \sim N(0,i)$ denote a Brownian motion on $[0,1]$, where $Z_0 = 0$. A Brownian bridge B_i is defined as

$$B_i \equiv Z_i - iZ_1. \tag{24}$$

Hence, this implies that $B_0 = B_1 = 0$ and B_i has continuous sample paths (a.s.). Equivalently, a Brownian bridge can be defined as a process distributed as a Brownian motion Z_i conditional on $Z_1 = 0$. As shown below, the variance of the Brownian bridge is larger for intermediate values of i. To obtain an identical distribution for all points $i \in [0,1)$, we define the *standardized shock* ϵ_i as follows

$$\epsilon_i \equiv \sqrt{12}\left(B_i - \int_0^1 B_j dj\right). \tag{25}$$

The following proposition summarizes the basic properties of B_i and ϵ_i.

Proposition 7. *Let B_i be a Brownian bridge and ϵ_i the standardized shock, then*

i. Brownian bridge:

$$\mathbb{E}[B_i] = 0, \qquad Var[B_i] = i(1-i), \qquad Cov(B_i, B_j) = \min\{i,j\} - ij. \tag{26}$$

ii. Standardized shock:

$$\mathbb{E}[\epsilon_i] = 0, \qquad Var[\epsilon_i] = 1, \qquad Cov(\epsilon_i, \epsilon_j) = 1 - 6|i-j|(1-|i-j|), \qquad Var\left[\int_0^1 \epsilon_i di\right] = 0. \tag{27}$$

Proof. **Brownian bridge**. The expected value is

$$\mathbb{E}[B_i] = \mathbb{E}[Z_i] - i\mathbb{E}[Z_1] = 0. \tag{28}$$

The variance is given by

$$\begin{aligned} Var[B_i] &= Var[Z_i(1-i) - i(Z_1 - Z_i)] \\ &= (1-i)^2 i + i^2(1-i) \\ &= i(1-i). \end{aligned} \tag{29}$$

Without loss of generality, suppose $0 \leq i \leq j \leq 1$, so $\min\{i, j\} = i$; then the covariance is

$$\begin{aligned} Cov(B_i, B_j) &= Cov(Z_i - iZ_1, Z_j - jZ_1) \\ &= Cov(Z_i, Z_j) - jCov(Z_i, Z_1) - iCov(Z_1, Z_j) + ijVar(Z_1) \\ &= i - ji - ij + ij \\ &= i - ij. \end{aligned} \tag{30}$$

Standardized shock. The expected value is

$$\mathbb{E}[\epsilon_i] = \sqrt{12}\left(\mathbb{E}[B_i] - \int_0^1 \mathbb{E}[B_j]dj\right) = 0. \tag{31}$$

The variance of the cross-sectional average of the Brownian bridge is

$$\begin{aligned} Var\left[\int_0^1 B_j dj\right] &= \int_0^1 \int_0^1 Cov(B_i, B_j) di dj \\ &= \int_0^1 \left[\int_0^j i(1-j)di + \int_j^1 j(1-i)di\right] dj \\ &= \int_0^1 \left[\frac{j^2}{2}(1-j) + j\left(1 - j - \frac{1-j^2}{2}\right)\right] dj \\ &= \frac{1}{2}\int_0^1 j(1-j)dj \\ &= \frac{1}{12}. \end{aligned} \tag{32}$$

The covariance between B_i and the cross-sectional average is

$$\begin{aligned} Cov(B_i, \int_0^1 B_j dj) &= \int_0^1 Cov(B_i, B_j)dj \\ &= \int_0^i j(1-i)dj + \int_i^1 i(1-j)dj \\ &= \frac{i^2}{2}(1-i) + i\left[1 - i - \frac{1-i^2}{2}\right] \\ &= \frac{i(1-i)}{2}. \end{aligned} \tag{33}$$

The variance of ϵ_i is then given by

$$\begin{aligned} Var[\epsilon_i] &= 12\left[Var(B_i) - 2Cov(B_i, \int_0^1 B_j dj) + Var\left[\int_0^1 B_j dj\right]\right] \\ &= 12\left(i(1-i) - i(1-i) + \frac{1}{12}\right) \\ &= 1. \end{aligned} \tag{34}$$

Suppose $0 \leq i \leq k \leq 1$. The covariance between ϵ_i and ϵ_k is given by

$$\begin{aligned}
Cov(\epsilon_i, \epsilon_k) &= 12Cov\left(B_i - \int_0^1 B_j dj, B_k - \int_0^1 B_j dj\right) \\
&= 12\left[Cov\left(B_i, B_k\right) - Cov\left(B_i, \int_0^1 B_j dj\right) - Cov\left(\int_0^1 B_j dj, B_k\right) + Var\left(\int_0^1 B_j dj\right)\right] \\
&= 12\left[\min\{i,k\} - ik - \frac{i(1-i)}{2} - \frac{k(1-k)}{2} + \frac{1}{12}\right] \\
&= 1 - 6i(k-i) - 6(k-i)(1-k) \\
&= 1 - 6(k-i)(1-(k-i)). \qquad (35)
\end{aligned}$$

Finally, we consider the variance of $\int_0^1 \epsilon_i di$:

$$\begin{aligned}
Var\left[\int_0^1 \epsilon_i di\right] &= \int_0^1 \int_0^1 Cov(\epsilon_i, \epsilon_j) didj \\
&= \int_0^1 \left[1 - 6\int_0^j (j - i - (j-i)^2) di - 6\int_j^1 (i - j - (i-j)^2) di\right] dj \\
&= \int_0^1 \left[1 - 6\left(j^2(1-j) - \frac{j^2}{2}(1-2j) - \frac{j^3}{3}\right) - 6\left(-(1-j)j(1+j) + \frac{1-j^2}{2}(1+2j) - \frac{1-j^3}{3}\right)\right] dj \\
&= \int_0^1 \left[1 - 6\left(\frac{1}{2} - \frac{1}{3}\right)\right] dj \\
&= 0. \qquad (36)
\end{aligned}$$

□

Because the vector $[B_{i_1}, B_{i_2}, \ldots, B_{i_K}]'$, for indices $i_1, \ldots, i_K \in [0,1]$, follows a multivariate normal distribution, similarly we deduce that $[\epsilon_{i_1}, \epsilon_{i_2}, \ldots, \epsilon_{i_K}]'$ follows a multivariate normal distribution with vector of means and variance-covariance matrix as described in the previous proposition.

A.2 Asymptotic analysis

We consider an approximation of the equilibrium around $\sigma_\theta = 0$. In particular, we solve for an approximation of the investors' consumption and portfolio decisions (C_0, C_s^i, Ω_j^i), firms' investment decisions (I^0, I^1), and return on assets $R_{s,j}^a$. More explicitly, for the variables without exposure to idiosyncratic risk, we consider the expansion

$$C_0 = C_0^* + \hat{C}_0 \sigma_\theta^2 + o(\sigma_\theta^2) \qquad (37)$$

$$I^k = I^{k,*} + \hat{I}^k \sigma_\theta^2 + o(\sigma_\theta^2), \qquad (38)$$

for $k = 0, 1$.

In the above expression, C_0^* and $I^{k,*}$ denote, respectively, the level of initial consumption and investment in technology k in the economy *without* idiosyncratic risk, i.e. $\sigma_\theta^2 = 0$. Our main interest lies in determining how these variables respond to the presence of idiosyncratic risk, i.e. to solve for the first-order impact of the idiosyncratic variance σ_θ^2 on these variables, given by the terms $\hat{C}_0$ and $\hat{I}^k$.

Regarding the variables exposed to idiosyncratic risk, their expansion in terms of σ_θ can be written as

$$C_s^i = C_s^* + \hat{C}_s\sigma_\theta^2 + \tilde{C}_s\bar{\epsilon}^{i,*}\sigma_\theta + o(\sigma_\theta^2) \tag{39}$$

$$R_{s,j}^a = R_s^{a,*} + \hat{R}_s^a\sigma_\theta^2 + \tilde{R}_s^a\epsilon_j\sigma_\theta + o(\sigma_\theta^2), \tag{40}$$

where $\bar{\epsilon}^{i,*}$ is an average over ϵ_j to be discussed below.

The term $\hat{R}_s^a$ now captures the impact of idiosyncratic risk on the *average* value of $R_{s,j}^a$, while $\tilde{R}_s^a$ captures the magnitude of idiosyncratic risk in $R_{s,j}^a$ in state s, i.e. $Var_s(R_{sj}^a) = (\tilde{R}_s^a)^2\sigma_\theta^2$, where the variance is conditional on the aggregate state s.

Finally, consider the expansion of the portfolio choice Ω_j^i

$$\Omega_j^i = \Omega_j^{i,*} + \hat{\Omega}_j^i\sigma_\theta^2 + o(\sigma_\theta^2). \tag{41}$$

Importantly, the portfolio choice is *not* determined when $\sigma_\theta = 0$, as the investor is indifferent regarding all firms in the participation set. Hence, $\Omega_j^{i,*}$ is not the solution when there is no idiosyncratic risk, insofar as the solution is indeterminate in that case, but the limit of the portfolio choice, Ω_j^i, as σ_θ^2 goes to zero. In contrast to C_s^* or C_0^*, for instance, which are considered as given when computing the perturbation coefficients, we need to solve for $\Omega_j^{i,*}$ jointly with the remaining perturbation coefficients.[39]

A.3 Ito-like formulas

Given the expansion for a variable, we may be interested in computing the expansion for functions of such a variable. For instance, given the coefficients in the expansions for C_s^i, we may want to compute the expansion for $c_s^i \equiv \log C_s^i$. The next lemma, a slight generalization of the Ito-like result discussed in Section 2, allows us to compute these expansions.

Lemma 2 (Ito-like)**.** *Let $F(\cdot)$ denote a twice-differentiable function and $X_{s,j} = X_s^* + \hat{X}_s\sigma_\theta^2 + \tilde{X}_s\epsilon_j\sigma_\theta$. Then,*

$$\mathbb{E}_s[F(X_{s,j}) - F(X_s^*)] = \left(F'(X_s^*)\hat{X}_s + \frac{1}{2}F''(X_s^*)\tilde{X}_s^2\right)\sigma_\theta^2 + o(\sigma_\theta^2), \qquad Var_s\left[F(\sigma_\theta\epsilon_j)\right] = F'(X_s^*)^2\tilde{X}_s^2\sigma_\theta^2 + o(\sigma_\theta^2). \tag{42}$$

and all the higher-order central moments are of order $o(\sigma_\theta^2)$.

Moreover, the covariance between a function of two shocks satisfies

$$Cov_s(F(X_{s,i}), F(X_{s,j})) = F'(X_s^*)^2\tilde{X}_s^2\sigma_\theta^2 Cov(\epsilon_i, \epsilon_j) + o(\sigma_\theta^2). \tag{43}$$

[39] Formally, the fact that the portfolio choice is indeterminate at σ_θ implies that the conditions on applying the implicit function theorem do not hold. The approximation procedure is then based on a *bifurcation theorem*; see Judd and Guu (2001) for a discussion of these issues.

Proof. Expanding $F(X_{s,j})$ in σ_θ, we obtain

$$\begin{aligned} F(X_{s,j}) &= F(X_s^*) + F'(X_s^*)\left(\hat{X}_s\sigma_\theta^2 + \tilde{X}_s\sigma_\theta\epsilon_j\right) + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\sigma_\theta^2\epsilon_j^2 + o(\sigma_\theta)^2 \\ &= F(X_s^*) + \left[F'(X_s^*)\hat{X}_s + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\right]\sigma_\theta^2 + F'(X_s^*)\tilde{X}_s\sigma_\theta\epsilon_j + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\sigma_\theta^2(\epsilon_j^2 - 1) + o(\sigma_\theta)^2. \end{aligned} \tag{44}$$

The expected value of the expression above is

$$\mathbb{E}_s[F(X_{s,j}) - F(X_s^*)] = \left[F'(X_s^*)\hat{X}_s + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\right]\sigma_\theta^2 + o(\sigma_\theta^2), \tag{45}$$

using $\mathbb{E}[\epsilon_j^2] = 1$.

The variance of $F(X_{s,j})$ is given by

$$\begin{aligned} Var_s[F(X_{s,j})] &= Var_s\left[F(X_{s,j}) - F(X_s^*) - \left(F'(X_s^*)\hat{X}_s + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\right)\sigma_\theta^2\right] \\ &= \mathbb{E}_s\left[\left(F'(X_s^*)\tilde{X}_s\sigma_\theta\epsilon_j + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\sigma_\theta^2(\epsilon_j^2 - 1)\right)^2\right] + o(\sigma_\theta^2) \\ &= F'(X_s^*)^2\tilde{X}_s^2\sigma_\theta^2 + o(\sigma_\theta^2). \end{aligned}$$

Consider now a central moment of order $k > 2$ of $F(\sigma_\theta\epsilon_j)$

$$\begin{aligned} \mathbb{E}_s\left[\left(F(X_{s,j}) - \mathbb{E}_s[F(X_{s,j})]\right)^k\right] &= \mathbb{E}\left[\sigma_\theta^k\left(F'(X_s^*)\tilde{X}_s + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\sigma_\theta(\epsilon_j^2 - 1)\right)^k\right] + o(\sigma_\theta^2) \\ &= o(\sigma_\theta^2). \end{aligned}$$

Finally, the covariance between $F(X_{s,i})$ and $F(X_{s,j})$ is given by

$$\begin{aligned} Cov_s(F(X_{s,i}), F(X_{s,j})) &= Cov_s\left(F'(X_s^*)\tilde{X}_s\sigma_\theta\epsilon_i + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\sigma_\theta^2(\epsilon_i^2 - 1), F'(X_s^*)\tilde{X}_s\sigma_\theta\epsilon_j + \frac{F''(X_s^*)}{2}\tilde{X}_s^2\sigma_\theta^2(\epsilon_j^2 - 1)\right) + o(\sigma_\theta^2) \\ &= F'(X_s^*)^2\tilde{X}_s^2\sigma_\theta^2 Cov\left(\epsilon_i, \epsilon_j\right) + o(\sigma_\theta^2). \end{aligned}$$

□

A corollary of the lemma above is that averaging a function of the shocks over $[0, 1)$ eliminates the idiosyncratic risk.

Corollary 2.

$$Var_s\left[\int_0^1 F(X_{s,j})dj\right] = 0. \tag{46}$$

Proof.

$$
\begin{aligned}
Var_s\left[\int_0^1 F(X_{s,j})dj\right] &= \int_0^1\int_0^1 Cov(F(X_{s,i}),F(X_{s,j}))didj \\
&= F'(X_s^*)^2\tilde{X}_s^2\sigma_\theta^2\int_0^1\int_0^1 Cov(\epsilon_i,\epsilon_j)didj \\
&= F'(X_s^*)^2\tilde{X}_s^2 Var\left[\int_0^1\epsilon_j dj\right] \\
&= 0, \qquad (47)
\end{aligned}
$$

up to the first-order in σ_θ^2. □

An implication of this result is that aggregate productivity equals Θ and it is not exposed to idiosyncratic risk. By choosing $F(X) = \Theta e^X$ and $X_j = -\frac{1}{2}\sigma_\theta^2 + \sigma_\theta\epsilon_j$, we deduce that $\theta_j = F(X_j)$, then $\mathbb{E}[\int_0^1\theta_j dj] = \Theta$ and $Var\left[\int_0^1\theta_j dj\right] = 0$, so $\int_0^1\theta_j dj = \Theta$, almost surely.

A.4 Portfolio choice: proof of Proposition 1

The next lemma characterizes the minimal-variance portfolio and shows that, in the small risk approximation, the optimal portfolio equals the minimal-variance portfolio.

Lemma 3. *Consider the portfolio problem (5). The portfolio Ω_j^i for investor in position i that minimizes $Var[\int_{0-}^1\epsilon_j d\Omega_j^i]$ subject to the participation constraint (4) is given by*[40]

$$
\Omega_j^i = \begin{cases} 0, & \text{if } j < i - 0.5\phi \\ \frac{1-\phi}{2} + j, & \text{if } i - 0.5\phi \le j < i + 0.5\phi \\ 1, & \text{if } j \ge i + 0.5\phi \end{cases}. \qquad (48)
$$

The variance under the optimal portfolio is

$$
Var\left[\int_{0-}^1 \epsilon_j d\Omega_j^i\right] = (1-\phi)^3. \qquad (49)
$$

Proof. To ease the notation, we focus on the case in which $i = 0.5\phi$, so the investor is allowed to choose among the assets in the interval $[0,\phi]$. We can obtain the solution to any other value of i by properly shifting the solution. The problem of minimizing the variance of the portfolio subject to the underdiversification constraint is given by

$$
\min_\Omega \frac{1}{2}\int_{0-}^{\phi}\int_{0-}^{\phi} Cov(\epsilon_i,\epsilon_j)d\Omega_i d\Omega_j, \qquad (50)
$$

[40]Following Gârleanu et al. (2015), we identify the index j with $j - \lfloor j \rfloor$, where $\lfloor x \rfloor$ is the largest integer weakly smaller than x, e.g. the indices -0.1 and 1.9 represent the same firm $j = 0.9$.

subject to

$$\int_{0-}^{\phi} d\Omega_i = 1. \tag{51}$$

The first-order condition is given by

$$\int_{0-}^{\phi} Cov(\epsilon_i, \epsilon_j) d\Omega_i = \lambda, \tag{52}$$

for all $j \in [0, \phi]$.

From the expression for the covariance of two shocks, we obtain

$$\begin{aligned}\int_{0-}^{\phi} Cov(\epsilon_i, \epsilon_j) d\Omega_i &= 1 - 6\int_{0-}^{j} \left[j - i - (j-i)^2\right] d\Omega_i - 6\int_{j}^{\phi} \left[i - j - (j-i)^2\right] d\Omega_i \\ &= 1 + 6\int_{0-}^{j} \Omega_i(-1 + 2(j-i))di - 6[(\phi - j) - (\phi - j)^2] + 6\int_{j}^{\phi} \Omega_i(1 + 2(j-i))di,\end{aligned}$$

where we applied integration by parts in the second equality and used $\Omega_\phi = 1$ and $\Omega_{0-} = 0$.

Because the expression above does not vary across assets j, its derivative with respect to j must be equal to zero

$$-\Omega_j + \int_{0-}^{\phi} \Omega_i di + \frac{1}{2} - \phi + j = 0. \tag{53}$$

Hence, we must have $\Omega_j = C_\Omega + j$ for some constant C_Ω. Plugging this functional form in the previous expression, we obtain

$$C_\Omega = \frac{1-\phi}{2}. \tag{54}$$

The Lagrange multiplier is then given by

$$\begin{aligned}\lambda &= 1 - 6C_\Omega j(1-j) - 3j^2(1-2j) - 4j^3 - 6[(\phi - j) - (\phi - j)^2] + 6C_\Omega\left[(\phi - j)(1+2j) - (\phi^2 - j^2)\right] \\ &\quad + 3(\phi^2 - j^2)(1+2j) - 12\frac{\phi^3 - j^3}{3} \\ &= 1 - 6\phi(1-\phi) + 6C_\Omega\phi(1-\phi) + 3\phi^2 - 4\phi^3 \\ &= 1 - 3\phi + 3\phi^2 - \phi^3 \\ &= (1-\phi)^3.\end{aligned} \tag{55}$$

The variance of the portfolio is given by

$$\int_{0-}^{\phi}\int_{0-}^{\phi} Cov(\epsilon_i, \epsilon_j) d\Omega_i d\Omega_j = \int_{0-}^{\phi} \lambda d\Omega_j = (1-\phi)^3. \tag{56}$$

□

Given the characterization of the minimal-variance portfolio, we can now prove proposition 1.

Proof. We begin by establishing the necessity of the Euler condition (7). Let Ω_j^i denote a (candidate) solution and consider the alternative $(1-\omega)\Omega_j^i + \omega\Omega_j^{j'}$, where $\Omega_j^{j'}$ is a cdf of a distribution giving all the weight to firm $j' \in \mathcal{P}^i$, i.e. $\Omega_j^{j'} = 0$ if $j < j'$ and $\Omega_j^{j'} = 1$ if $j \geq j'$. If Ω_j^i is optimal, then the derivative of the objective function with respect to ω evaluated at $\omega = 0$ is zero:

$$\mathbb{E}\left[u'(C_s^i)\left(\int_{0-}^1 R_s^a(\theta_j)d\Omega_j^{j'} - \int_{0-}^1 R_s^a(\theta_j)d\Omega_j^i\right)K_s\right] = 0. \tag{57}$$

Rearranging the expression above, we obtain

$$\mathbb{E}\left[u'(C_s^i)R_s^a(\theta_{j'})K_s\right] = \mathbb{E}\left[u'(C_s^i)\int_{0-}^1 R_s^a(\theta_j)d\Omega_j^i K_s\right]. \tag{58}$$

Combining the expression above with (6), we obtain (7). We now derive an asymptotic expansion of the expression above. Consumption at state s of investor i is given by

$$C_s^i = K_s \int_{0-}^1 R_s^a(\theta_j)d\Omega_j^i. \tag{59}$$

We can write capital as follows

$$K_s = K_s^* + \hat{K}_s\sigma_\theta^2 + o(\sigma_\theta^2), \tag{60}$$

where

$$K_s^* \equiv (1 + \chi^*\varphi_s^e)I^*$$
$$\hat{K}_s \equiv \varphi_s^e I^* \hat{\chi} + (1 + \chi^*\varphi_s^e)\hat{I}.$$

The return on assets for firm j as

$$R_{s,j}^a = R_s^{a,*} + \alpha\Theta^\alpha(K_s^*)^{\alpha-1}\epsilon_j\sigma_\theta - \alpha(1-\alpha)\Theta^\alpha(K_s^*)^{\alpha-1}\frac{\hat{K}_s}{K_s^*}\sigma_\theta^2 + o(\sigma_\theta^2), \tag{61}$$

where $R_s^{a,*} = 1 - \delta + \alpha\Theta^\alpha(K_s^*)^{\alpha-1}$.

Expanding the average return on assets for the firms in the portfolio of investor i, we obtain

$$\int_{0-}^1 R_{s,j}^a d\Omega_j^i = R_s^{a,*} + \alpha\Theta^\alpha(K_s^*)^{\alpha-1}\epsilon^{i,*}\sigma_\theta - \alpha(1-\alpha)\Theta^\alpha(K_s^*)^{\alpha-1}\frac{\hat{K}_s}{K_s^*}\sigma_\theta^2 + o(\sigma_\theta^2), \tag{62}$$

where $\epsilon^{i,*} = \int_{0-}^1 \epsilon_j d\Omega_j^{i,*}$.

The asymptotic expansion of consumption of investor i in state s is

$$C_s^i = C_s^* + \hat{C}_s\sigma_\theta^2 + \tilde{C}_s\epsilon^{i,*}\sigma_\theta + o(\sigma_\theta^2), \tag{63}$$

where

$$
\begin{aligned}
C_s^* &= R_s^{a,*} K_s^* \\
\hat{C}_s &= -\alpha(1-\alpha)\Theta^\alpha (K_s^*)^{\alpha-1}\hat{K}_s + R_s^{a,*}\hat{K}_s \\
\tilde{C}_s &= \alpha(\Theta K_s^*)^\alpha.
\end{aligned}
$$

Computing the expansion of the marginal utility of consumption, we obtain

$$
u'(C_s^i) = (C_s^*)^{-\gamma} - \gamma (C_s^*)^{-\gamma-1}\alpha(\Theta K_s^*)^\alpha \epsilon^{i,*}\sigma_\theta + \mathcal{O}(\sigma_\theta^2). \tag{64}
$$

The first-order condition with respect to the portfolio choice can be written as

$$
\mathbb{E}\left[u'(C_s^i)(R_s(\theta_j) - R_s(\theta_{j'}))K_s\right] = 0, \tag{65}
$$

for any two firms j and j' in the participation set.

Expanding the expression above, we obtain

$$
\mathbb{E}\left[-\gamma(C_s^*)^{-\gamma}\frac{\alpha\Theta^\alpha(K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha(K_s^*)^{\alpha-1}}\epsilon^{i,*}\sigma_\theta\alpha\Theta^\alpha(K_s^*)^{\alpha-1}(\epsilon_j - \epsilon_{j'})\sigma_\theta K_s^*\right] + o(\sigma_\theta^2) = 0. \tag{66}
$$

Rearranging the expression above, we obtain

$$
\mathbb{E}\left[\epsilon^{i,*}\epsilon_j\right] = \mathbb{E}\left[\epsilon^{i,*}\epsilon_{j'}\right] \Rightarrow Cov(\epsilon^{i,*}, \epsilon_j) = Cov(\epsilon^{i,*}, \epsilon_{j'}). \tag{67}
$$

Hence, all assets in the participation set have the same covariance with the payoff $\Omega_j^{i,*}$, which implies that $\Omega_j^{i,*}$ is the minimum-variance portfolio. From lemma 3, we deduce that $Var[\int_{0-}^1 \epsilon_j d\Omega_j^{*,i}] = (1-\phi)^3$. Since the minimal-variance portfolio equalizes the covariance of the portfolio with any asset in the participation set, we have that $Cov(\epsilon^{i,*}, \epsilon_j) = (1-\phi)^3$ for any $j \in \mathcal{P}^i$. □

A.5 Idiosyncratic risk premium

Define the log stochastic discount factor for investor i

$$
m_s^i \equiv \log\beta - \gamma(c_s^i - c_0).
$$

Log-consumption can be written as

$$
c_s^i = c_s^* + \frac{\alpha\Theta^\alpha(K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha(K_s^*)^{\alpha-1}}\epsilon^{i,*}\sigma_\theta + \mathcal{O}(\sigma_\theta^2). \tag{68}
$$

The (log) risk-free interest rate $r_f \equiv \log R_f$ satisfies

$$\begin{aligned}1 &= \mathbb{E}\left[e^{m_s^i + r_f}\right] \\ &\approx 1 + \mathbb{E}[m_s^i] + \frac{1}{2}Var[m_s^i] + r_f.\end{aligned}$$

This implies that

$$r_f = -\mathbb{E}[m_s^i] - \frac{1}{2}Var[m_s^i]. \tag{69}$$

Notice that, because c_s^i is identically distributed across investors i, the risk-free rate displayed above does not depend on i.

We now consider the expected return on the firms. In a symmetric equilibrium, we have that $P_j = \sum_{k=0}^{1} I^k$. Then, the return on firm j is given by

$$R_{s,j} \equiv \frac{R_{s,j}^a K_s}{P} = R_{s,j}^a \frac{\sum_{k=0}^{1} \varphi_s^k I^k}{\sum_{k=0}^{1} I^k} = R_{s,j}^a(1 + \varphi_s^e \chi), \tag{70}$$

where $\chi \equiv \frac{I^1}{I^0 + I^1}$ is the share invested in the risky technology and $\varphi_s^e = \varphi_s^1 - \varphi_s^0$ is the excess payoff of the risky technology.

The log return on firm j is then defined as

$$r_{s,j} = r_{s,j}^a + r_s^I, \tag{71}$$

where $r_{s,j} \equiv \log R_{s,j}$, $r_{s,j}^a \equiv \log R_{s,j}^a$, and $r_s^I \equiv \log(1 + \varphi_s^e \chi)$.

We can write the excess return for firm j as follows

$$r_{s,j} = r_s^* + \frac{\alpha \Theta^\alpha (K_s^*)^{\alpha-1}}{1 - \delta + \alpha \Theta^\alpha (K_s^*)^{\alpha-1}} \epsilon_j \sigma_\theta + \mathcal{O}(\sigma_\theta^2). \tag{72}$$

The pricing equation for firm j can be written as

$$\begin{aligned}1 &= \mathbb{E}[e^{m_s^i + r_{s,j}}] \\ &\approx 1 + \mathbb{E}[m_s^i + r_{s,j}] + \frac{1}{2}Var[m_s^i + r_{s,j}],\end{aligned} \tag{73}$$

where firm j belongs to the participation set of firm i.

Rearranging the expression above, we obtain

$$\mathbb{E}[r_j] - r_f + \frac{1}{2}Var[r_j] = \gamma\left[\mathbb{E}\left[Cov_s(c^i, r_j)\right] + Cov(\bar{c}^i, \bar{r}_j)\right], \tag{74}$$

where we applied the law of total covariance, used the definition of the log stochastic discount factor m_s^i, and defined $\bar{c}_s^i \equiv \mathbb{E}_s[c_s^i]$ and $\bar{r}_{s,j} \equiv \mathbb{E}_s[r_j]$ as the expectation over the idiosyncratic state, conditional on the aggregate state s for consumption and excess return.

We can then write the previous expression as

$$\mathbb{E}[r_j] - r_f + \frac{1}{2}Var[r_j] = \gamma Cov(\bar{c}, \bar{r}_j) + \gamma \mathbb{E}\left[\left(\frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}\sigma_\theta\right)^2 Cov(\epsilon^{i,*}, \epsilon_j)\right]. \tag{75}$$

From the properties of the optimal portfolio allocation, we deduce that

$$Cov(\epsilon^{i,*}, \epsilon_j) = (1-\phi)^3. \tag{76}$$

Finally, we obtain the expression for the excess return on firm j

$$\mathbb{E}[r_j^e] + \frac{1}{2}Var[r_j^e] = \gamma Cov(\bar{c}, \bar{r}_j^e) + \gamma(1-\phi)^3 \mathbb{E}\left[\sigma_s^2\right]. \tag{77}$$

where $\sigma_s = \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}\sigma_\theta$.

A.6 Investment and aggregate risk-taking: proof of Proposition 2

Proof. Using the fact that $P = I$, then we can write the investor's Euler equation as follows

$$1 = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0^i)} R_s^a(\theta_j)(1+\chi\varphi_s^e)\right], \tag{78}$$

where $j \in \mathcal{P}^i$.

The investment Euler equation for firm j is

$$1 = \mathbb{E}\left[M_s^j R_{s,j}^a \varphi_s^k\right], \tag{79}$$

for $k = 0, 1$.

The stochastic discount factor for firm j is given by

$$M_{s,j} = \int_0^1 \beta \frac{u'(C_s^i)}{u'(C_0)} dF_{i,j}, \tag{80}$$

given a cdf satisfying $\int_{\{i:d(i,j)\leq 0.5\phi\}} dF_{i,j}$.

We first assume that the F_i^j gives all the weight to a single investor i and then show that the identity of this investor does not matter, so we obtain the same results for any distribution F_i^j. Under this assumption, we obtain the following Euler condition

$$0 = \mathbb{E}\left[u'(C_s^i) R_{s,j}^a \varphi_s^e\right]. \tag{81}$$

We first consider an expansion of the marginal utility

$$u'(C_s^i) = (C_s^*)^{-\gamma}\left[1 - \gamma\left(\frac{\hat{C}_s}{C_s^*}\sigma_\theta^2 + \frac{\tilde{C}_s}{C_s^*}\epsilon^{i,*}\sigma_\theta\right) + \frac{\gamma(\gamma+1)}{2}\left(\frac{\tilde{C}_s}{C_s^*}\right)^2 \phi_u \sigma_\theta^2\right] + o(\sigma_\theta^2), \tag{82}$$

and the return on firm j

$$R^a_{s,j} = R^{a,*}_s \left[1 + \frac{\hat{R}^a_s}{R^{a,*}_s}\sigma^2_\theta + \frac{\tilde{R}^a_s}{R^{a,*}_s}\epsilon_j\sigma_\theta\right] + o(\sigma^2_\theta). \tag{83}$$

Plugging the previous two expressions into the Euler equation, we obtain

$$0 = \mathbb{E}\left[(C^*_s)^{-\gamma} R^{a,*}_s \varphi^e_s \left(-\gamma\frac{\hat{C}_s}{C^*_s} + \frac{\hat{R}^a_s}{R^{a,*}_s} + \frac{\gamma(\gamma+1)}{2}\left(\frac{\tilde{C}_s}{C^*_s}\right)^2 \phi_u - \gamma\frac{\tilde{C}_s}{C^*_s}\frac{\tilde{R}^a_s}{R^{a,*}_s}\phi_u\right)\right]\sigma^2_\theta + o(\sigma^2_\theta). \tag{84}$$

Notice that the above expression does not depend on i, so we would obtain the same expression for any cdf $F_{i,j}$. Consider now the Euler equation for the safe technology

$$u'(C_0) = \mathbb{E}\left[\beta u'(C^i_s) R^a_{s,j}\right]. \tag{85}$$

Expanding the expression above, we obtain

$$-\gamma\frac{\hat{C}_0}{C^*_0}\sigma^2_\theta = \mathbb{E}\left[\beta\left(\frac{C^*_s}{C^*_0}\right)^{-\gamma} R^{a,*}_s \left(-\gamma\frac{\hat{C}_s}{C^*_s} + \frac{\hat{R}^a_s}{R^{a,*}_s} + \frac{\gamma(\gamma+1)}{2}\left(\frac{\tilde{C}_s}{C^*_s}\right)^2 \phi_u - \gamma\frac{\tilde{C}_s}{C^*_s}\frac{\tilde{R}^a_s}{R^{a,*}_s}\phi_u\right)\right]\sigma^2_\theta + o(\sigma^2_\theta). \tag{86}$$

where

$$\begin{aligned} \hat{C}^*_0 &= E_0 - I^* \\ \hat{C}_0 &= -\hat{I}. \end{aligned} \tag{87}$$

We can write the above coefficients in terms of only $\hat{\chi}$ and $\hat{I}$

$$\begin{aligned} \frac{\hat{C}_s}{C^*_s} &= \left[1 - (1-\alpha)\frac{\alpha\Theta^\alpha (K^*_s)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha(K^*_s)^{\alpha-1}}\right]\frac{\hat{K}_s}{K^*_s} \\ \frac{\hat{R}^a_s}{R^{a,*}_s} &= -(1-\alpha)\frac{\alpha\Theta^\alpha (K^*_s)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha(K^*_s)^{\alpha-1}}\frac{\hat{K}_s}{K^*_s} \\ \frac{\hat{K}_s}{K^*_s} &= \frac{\varphi^e_s}{1+\chi^*\varphi^e_s}\hat{\chi} + \hat{I} \\ \frac{\hat{C}_0}{C^*_0} &= -\frac{I^*}{E_0 - I^*}\frac{\hat{I}}{I^*}. \end{aligned} \tag{88}$$

Combining the expressions above, we obtain the system

$$\begin{bmatrix} a_{\chi\chi} & a_{\chi I} \\ a_{I\chi} & a_{II} \end{bmatrix}\begin{bmatrix} \hat{\chi} \\ \hat{I} \end{bmatrix} = \begin{bmatrix} b_\chi \\ b_I \end{bmatrix}, \tag{89}$$

where

$$
\begin{aligned}
a_{\chi\chi} &= Cov\left((C_s^*)^{-\gamma} R_s^{a,*} \varphi_s^e, \left[\gamma \left(1 - (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right) + (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right] \frac{\varphi_s^e}{1+\chi^* \varphi_s^e} \right) \\
a_{\chi I} &= Cov\left((C_s^*)^{-\gamma} R_s^{a,*} \varphi_s^e, \gamma \left(1 - (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right) + (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right) \\
a_{I\chi} &= \mathbb{E}\left[\beta \left(\frac{C_s^*}{C_0^*} \right)^{-\gamma} R_s^{a,*} \left[\gamma \left(1 - (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right) + (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right] \frac{\varphi_s^e}{1+\chi^* \varphi_s^e} \right] \\
a_{II} &= \mathbb{E}\left[\beta \left(\frac{C_s^*}{C_0^*} \right)^{-\gamma} R_s^{a,*} \left[\gamma \left(1 - (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right) + (1-\alpha) \frac{\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{1-\delta+\alpha\Theta^\alpha (K_s^*)^{\alpha-1}} \right] \right] + \frac{\gamma}{E_0 - I^*},
\end{aligned}
\tag{90}
$$

and

$$
\begin{aligned}
b_\chi &= Cov\left((C_s^*)^{-\gamma} R_s^{a,*} \varphi_s^e, \frac{\gamma(\gamma-1)}{2} \left(\frac{\tilde{R}_s^a}{R_s^*} \right)^2 \right) \phi_u \\
b_I &= \mathbb{E}\left[\beta \left(\frac{C_s^*}{C_0^*} \right)^{-\gamma} R_s^{a,*} \frac{\gamma(\gamma-1)}{2} \left(\frac{\tilde{R}_s^a}{R_s^*} \right)^2 \right] \phi_u.
\end{aligned}
$$

Assuming $\gamma > 1$, we can show that $a_{\chi\chi} > 0$, $a_{\chi I} > 0$, $a_{I\chi} < 0$, $a_{II} > 0$.[41] The remaining coefficients satisfy $b_\chi < 0$ and $b_I > 0$. Solving the system above, we deduce that

$$
\begin{aligned}
\hat{\chi} &= \frac{a_{II} b_\chi - a_{\chi I} b_I}{a_{\chi\chi} a_{II} - a_{\chi I} a I\chi} \\
\hat{I} &= \frac{a_{\chi\chi} b_I - a_{I\chi} b_\chi}{a_{\chi\chi} a_{II} - a_{\chi I} a_{I\chi}}.
\end{aligned}
\tag{91}
$$

The response of aggregate risk-taking satisfies $\hat{\chi} < 0$, but the sign of the coefficient $\hat{I}$ is ambiguous. Suppose now that $\hat{\chi} = 0$. The solution in this case can be obtained by simply setting $a_{I\chi} = 0$ in the expression above for $\hat{I}$:

$$
\hat{I} = \frac{b_I}{a_{II}} > 0. \tag{92}
$$

□

[41] For $a_{I\chi}$, write $a_{I\chi} = Cov\left(\beta \left(\frac{C_s^*}{C_0^*} \right)^{-\gamma} R_s^{a,*} \varphi_s^e, \frac{\gamma(1-\delta)+(\gamma\alpha+1-\alpha)\alpha\Theta^\alpha (K_s^*)^{\alpha-1}}{(1-\delta)K_s^* + \alpha\Theta^\alpha (K_s^*)^\alpha} \right) I^* <0$, where we used the fact that $\mathbb{E}\left[(C_s^*)^{-\gamma} R_s^{a,*} \right] = 0$ to write the expression as a covariance.

Pricing kernel for aggregate payoffs

We now derive expression (11). From the expansion for the marginal utility of consumption in period 1 (82) and for ROA (83), we obtain

$$\begin{aligned}\mathbb{E}_s\left[(C_s^i)^{-\gamma}R_{s,j}^a\right] &= (C_s^*)^{-\gamma}R_s^{a,*} + (C_s^*)^{-\gamma}R_s^{a,*}\left(-\gamma\frac{\hat{C}_s}{C_s^*} + \frac{\hat{R}_s^a}{R_s^{a,*}} + \frac{\gamma(\gamma+1)}{2}\left(\frac{\tilde{C}_s}{C_s^*}\right)^2\phi_u - \gamma\frac{\tilde{C}_s}{C_s^*}\frac{\tilde{R}_s^a}{R_s^{a,*}}\phi_u\right)\sigma_\theta^2 + o(\sigma_\theta^2) \\ &= \overline{C}_s^{-\gamma}\overline{R}_s^a + (C_s^*)^{-\gamma}R_s^{a,*}\frac{\gamma(\gamma-1)}{2}\phi_u\sigma_s^2 + o(\sigma_\theta^2), \end{aligned} \tag{93}$$

using $\overline{C}_s^{-\gamma}\overline{R}_s^a = (C_s^*)^{-\gamma}R_s^{a,*} + (C_s^*)^{-\gamma}R_s^{a,*}\left(-\gamma\frac{\hat{C}_s}{C_s^*} + \frac{\hat{R}_s^a}{R_s^{a,*}}\right)\sigma_\theta^2 + o(\sigma_\theta^2)$ and $\sigma_s^2 = \frac{\tilde{C}_s}{C_s^*}\frac{\tilde{R}_s^a}{R_s^{a,*}}\sigma_\theta^2 = \left(\frac{\tilde{C}_s}{C_s^*}\right)^2\sigma_\theta^2$.

Finally, using $(C_s^*)^{-\gamma}R_s^{a,*}\sigma_\theta^2 = \overline{C}_s^{-\gamma}\overline{R}_s^a\sigma_\theta^2 + o(\sigma_\theta^2)$, we obtain

$$\mathbb{E}_s\left[(C_s^i)^{-\gamma}R_{s,j}^a\right] \approx \overline{C}_s^{-\gamma}\overline{R}_s^a \times \exp\left(\frac{\gamma(\gamma-1)}{2}\phi_u\sigma_s^2\right). \tag{94}$$

B Appendix for Section 3

For a given pair (κ_0, κ_1), the derivative of V with respect to Δ, at $\Delta = 0$, is

$$\begin{aligned}V'(0) &= -u'(C_0)(\kappa_0+\kappa_1) + \beta\mathbb{E}[u'(C_s^i)R_s^{a,i}(\kappa_0+\kappa_1\varphi_s^1)] + \beta\mathbb{E}\left[u'(C_s^i)K_s\left(\frac{\partial R_s^{a,i}}{\partial\Delta} + \frac{\partial\tau_s(\Delta)}{\partial\Delta}\right)\right] \\ &= \beta\mathbb{E}\left[u'(C_s^i)K_s\left(\frac{\partial R_s^{a,i}}{\partial\Delta} + \frac{\partial\tau_s(\Delta)}{\partial\Delta}\right)\right],\end{aligned} \tag{95}$$

where the second equality uses the Euler equations for investors and firms.

We now compute the derivative of the ROA for firm j with respect to Δ

$$\left.\frac{\partial R_{s,j}^a(\Delta)}{\partial\Delta}\right|_{\Delta=0} = -(1-\alpha)\alpha\theta_j\Theta^{\alpha-1}K_s^{\alpha-2}(\kappa_0+\kappa_1\varphi_s^1). \tag{96}$$

The derivative of the tax rate is given by

$$\left.\frac{\partial\tau_s(\Delta)}{\partial\Delta}\right|_{\Delta=0} = \alpha(1-\alpha)\Theta^\alpha K_s^{\alpha-2}(\kappa_0+\kappa_1\varphi_s^1). \tag{97}$$

Taking averages over the portfolio of the first expression and combining these averages with the second, we obtain

$$\left.\left(\frac{\partial R_s^{a,i}}{\partial\Delta} + \frac{\partial\tau_s(\Delta)}{\partial\Delta}\right)\right|_{\Delta=0} = -(1-\alpha)\alpha\left(\theta^i - \Theta\right)\Theta^{\alpha-1}K_s^{\alpha-2}(\kappa_0+\kappa_1\varphi_s^1). \tag{98}$$

Plugging the expression above into the expression for $V'(0)$, we obtain

$$V'(0) = -(1-\alpha)\beta\mathbb{E}\left[u'(C_s^i)\left(R_s^{a,i} - \overline{R}_s^a\right)(\kappa_0 + \kappa_1\varphi_s^1)\right]$$
$$= -(1-\alpha)\beta\mathbb{E}\left[Cov_s(u'(C_s^i), R_s^{a,i})(\kappa_0 + \kappa_1\varphi_s^1)\right]. \tag{99}$$

The covariance above can be written as

$$Cov_s(u'(C_s^i), R_s^{a,i}) = Cov_s\left(e^{-\gamma r_s^{a,i}}, e^{r_s^{a,i}}\right)K_s^{-\gamma}$$
$$= -\gamma\phi_u\sigma_s^2(C_s^*)^{-\gamma}\overline{R}_s^{a,*} + o(\sigma_\theta^2). \tag{100}$$

The derivative of the value function can then be written as

$$\frac{V'(0)}{u'(C_0)} = (1-\alpha)\gamma\phi_u\mathbb{E}\left[\beta\frac{u'(C_s^*)}{u'(C_0)}R_s^{a,*}\sigma_s^2(\kappa_0 + \kappa_1\varphi_s^1)\right] + o(\sigma_\theta^2). \tag{101}$$

Up to the first-order in σ_θ^2, we can write

$$\frac{V'(0)}{u'(C_0)} = (1-\alpha)\gamma\phi_u\mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}\sigma_s^2(\kappa_0 + \kappa_1\varphi_s^1)\right] + o(\sigma_\theta^2). \tag{102}$$

From the Euler condition for the riskless technology, we deduce that

$$\mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}\right] = 1. \tag{103}$$

Finally, define the *risk-neutral probabilities* as follows[42]

$$\mathbb{E}^{\mathbb{Q}}[X_s] \equiv \mathbb{E}\left[\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}X_s\right], \tag{104}$$

for any random variable X_s.

This allows us to write

$$\frac{V'(0)}{u'(C_0)} = (1-\alpha)\gamma\phi_u\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2(\kappa_0 + \kappa_1\varphi_s^1)\right] + o(\sigma_\theta^2). \tag{105}$$

We can use these equations to derive both propositions and the corollary, as below.

Proof of Proposition 3. Take $\kappa_0 = 1$ and $\kappa_1 = 0$, then $\frac{V'(0)}{u'(c_0)} = (1-\alpha)\gamma\phi_u\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2\right]$.

[42]Note that $\beta\frac{u'(C_s^i)}{u'(C_0)}R_s^{a,i}$ is the relevant pricing kernel for payoffs in terms of capital in period, i.e. before production takes place. Since the expectation of this pricing kernel is one, there is no risk-free rate dividing the expression.

We now show that the idiosyncratic variance risk premium is positive.

$$\mathbb{E}^{\mathbb{Q}}[\sigma_s^2] - \mathbb{E}[\sigma_s^2] = Cov\left(\beta \frac{u'(C_s^i)}{u'(C_0)} R_s^{a,i}, \sigma_s^2\right)$$
$$= Cov\left(\beta \frac{(C_s^i)^{-(\gamma-1)}}{C_0^{-\gamma}} \frac{1}{K_s}, \sigma_s^2\right) > 0, \tag{106}$$

where the inequality follows from C_s^i being increasing in K_s, σ_s^2 being decreasing in K_s, and the assumption $\gamma \geq 1$. □

Proof of Corollary 1. Immediately following from (105) after imposing $\alpha = 1$. □

Proof of Proposition 4. Taking $\kappa_0 = \frac{\mathbb{E}^{\mathbb{Q}}[\varphi^1]}{\sqrt{Var^{\mathbb{Q}}[\varphi^1]}}$ and $\kappa_1 = -\frac{1}{\sqrt{Var^{\mathbb{Q}}[\varphi^1]}}$, we have

$$\frac{V'(0)}{u'(C_0)} = -(1-\alpha)\gamma\phi_u \frac{\mathbb{E}^{\mathbb{Q}}\left[\sigma_s^2\left(\varphi_s^1 - \mathbb{E}^{\mathbb{Q}}[\varphi^1]\right)\right]}{\sqrt{Var^{\mathbb{Q}}[\varphi^1]}}$$
$$= -(1-\alpha)\gamma\phi_u \frac{Cov^{\mathbb{Q}}\left(\sigma_s^2, \varphi_s^1\right)}{\sqrt{Var^{\mathbb{Q}}[\varphi^1]}},$$

where we used the fact that $\mathbb{E}^{\mathbb{Q}}[\varphi_s^1] = 1$. We can also rewrite the first line as

$$\frac{V'(0)}{u'(C_0)} = -(1-\alpha)\gamma\phi_u \frac{1}{\sqrt{q_h q_l}\left(\varphi_h^1 - \varphi_l^1\right)}\left[\left(q_h q_l \sigma_h^2\left(\varphi_h^1 - \varphi_l^1\right) - q_h q_l \sigma_l^2\left(\varphi_h^1 - \varphi_l^1\right)\right)\right]$$
$$= (1-\alpha)\gamma\phi_u\sqrt{q_h q_l}\left(\sigma_l^2 - \sigma_h^2\right).$$

where the probabilities are to be interpreted as risk-neutral probabilities.

The idiosyncratic variance risk premium can be written as

$$\mathbb{E}^{\mathbb{Q}}[\sigma_s^2] - \mathbb{E}[\sigma_s^2] = (q_l - p_l)(\sigma_l^2 - \sigma_h^2) \tag{107}$$

Combining the previous two equations, we obtain the expression in the proposition. □

B.1 Extensions

B.1.1 Intermediate goods

Environment. We now consider an environment in which final goods are produced using capital and intermediate goods as inputs. For simplicity, labor is no longer a factor of production. In place of workers consuming the labor share, there are intermediate-goods entrepreneurs who consume their profits in period 1. The production of final goods is given by the production function $(\theta_j K_{s,j})^{\alpha} X_{s,j}^{1-\alpha}$, where $X_{s,j}$ denotes the use of intermediate goods by firm j in state s. Let Q denote the price of intermediate goods, then we obtain expressions for the demand for intermediates and final goods producers profits that are analogous to the ones with labor input:

$$X_{s,j} = \left[\frac{1-\alpha}{Q_s}\right]^{\frac{1}{\alpha}} \theta_j K_{s,j}, \qquad \pi_{s,j} = \alpha\theta_j \left[\frac{1-\alpha}{Q_s}\right]^{\frac{1-\alpha}{\alpha}}. \tag{108}$$

Intermediate goods are produced using a decreasing returns to scale technology. In particular, to produce X_s units of the intermediate good, $\frac{X_s^{1+\phi}}{1+\phi}$ units of the final good are needed, where $\phi > 0$. The problem of the intermediate-goods producer is

$$\pi_s^X = \max_{X_s} \left[Q_s X_s - \frac{X_s^{1+\phi}}{1+\phi}\right]. \tag{109}$$

The first-order condition for this problem is

$$Q_s = X_s^{\phi}. \tag{110}$$

Notice that if $\phi = 0$ the price of intermediate goods is fixed, while the quantity of intermediate goods is fixed if $\phi \to \infty$. The market clearing condition for intermediate goods is given by $\int_0^1 X_{s,j} dj = X_s$. Plugging the demand and supply for intermediate goods into the market clearing condition, we obtain

$$X_s = (1-\alpha)^{\frac{1}{\alpha+\phi}} (\Theta K_s)^{\frac{\alpha}{\alpha+\phi}}, \qquad Q_s = (1-\alpha)^{\frac{\phi}{\alpha+\phi}} (\Theta K_s)^{\frac{\alpha\phi}{\alpha+\phi}}. \tag{111}$$

The profit of intermediate-goods producers is given by

$$\pi_s^X = \frac{\phi}{1+\phi}(1-\alpha)^{\frac{1+\phi}{\alpha+\phi}} (\Theta K_s)^{\frac{\alpha(1+\phi)}{\alpha+\phi}}. \tag{112}$$

The ROA of a final-goods producer can be written as

$$R_{s,j}^a = 1 - \delta + \alpha\theta_j (1-\alpha)^{\frac{1-\alpha}{\alpha+\phi}} (\Theta K_s)^{\frac{(\alpha-1)\phi}{\alpha+\phi}}. \tag{113}$$

Notice that as $\phi \to \infty$, we recover the expression we obtained for the case with labor. We assume that a fraction ω_I^X of the profits of the intermediate-goods sector goes to investors and the fraction $1 - \omega_I^X$ remains with intermediate-goods entrepreneurs. The setting with $\phi \to \infty$ and $\omega_I^X = 0$ basically corresponds to the one discussed in the main text. The main distinction between this setup and the baseline model is that the variable input has a positive elasticity, whether it is an intermediate input or labor is not crucial for our results.

Idiosyncratic risk externalities. Consider the impact on the welfare of investors of a perturbation on investment

$$V(\Delta) = \max_{\Omega_j^i} \left\{ u\left(E_0 - \sum_{k=0}^{1} I^k(\Delta)\right) + \beta\mathbb{E}\left[u\left(\int_{0-}^{1} R_{s,j}^a(\Delta) d\Omega_j^i K_s + \omega_I^X \pi_s^X + T_s\right)\right]\right\}, \tag{114}$$

where

$$T_s = (1-\omega_I^X)\pi_s^X - C_s^X, \tag{115}$$

and C_s^X denotes the consumption of intermediate-goods entrepreneurs in laissez-faire.

The derivative of $V(\Delta)$ is given by

$$V'(0) = \beta\mathbb{E}\left[u'(C_s^i)\left(\frac{\partial R_s^{a,i}}{\partial \Delta}K_s + \frac{\partial \pi_s^X}{\partial \Delta}\right)\right]. \tag{116}$$

The derivative of the ROA and intermediate-goods profits with respect to Δ

$$\left.\frac{\partial R_{s,j}^a}{\partial \Delta}K_s\right|_{\Delta=0} = -\frac{\phi(1-\alpha)^{\frac{1+\phi}{\alpha+\phi}}}{\alpha+\phi}\alpha\theta_j\left(\Theta K_s\right)^{\frac{(\alpha-1)\phi}{\alpha+\phi}}\left(\kappa_0+\kappa_1\varphi_s^1\right) \tag{117}$$

$$\left.\frac{\partial \pi_s^X}{\partial \Delta}\right|_{\Delta=0} = \frac{\phi(1-\alpha)^{\frac{1+\phi}{\alpha+\phi}}}{\alpha+\phi}\alpha\Theta\left(\Theta K_s\right)^{\frac{(\alpha-1)\phi}{\alpha+\phi}}\left(\kappa_0+\kappa_1\varphi_s^1\right). \tag{118}$$

Hence, we can write the derivative of the value function as

$$V'(0) = -\frac{\phi(1-\alpha)}{\alpha+\phi}\beta\mathbb{E}\left[Cov_s(u'(C_s^i), R_s^{a,i})(\kappa_0+\kappa_1\varphi_s^1)\right]. \tag{119}$$

The expression above is analogous to the one we derived in the case with labor. The only difference is the constant of proportionality which is not $1-\alpha$ but instead $\frac{\phi(1-\alpha)}{\alpha+\phi}$. Hence, allowing for an elastic response of the variable input dampens the effect. For instance, if we set $\phi = 1$ and $\alpha = 0.3$, this implies a reduction in the effect in the order of 20%.

B.1.2 CES production function

We now assume that capital and labor are combined according to a CES production function. The problem of a firm in period 1 is then given by

$$\max_L \left[\alpha(\theta_j K_{s,j})^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)L^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}} - W_s L. \tag{120}$$

The demand for labor is given by

$$W_s = (1-\alpha)\left[\alpha\left(\frac{\theta_j K_{s,j}}{L_{s,j}}\right)^{\frac{\epsilon-1}{\epsilon}} + 1 - \alpha\right]^{\frac{1}{\epsilon-1}}. \tag{121}$$

The (effective) capital-labor is then equalized across firms. Profit per unit of capital for firm j can be written as

$$\pi_{s,j} = \alpha\theta_j\frac{W_s(\Theta K_s)^{-\frac{1}{\epsilon}}}{1-\alpha}. \tag{122}$$

The wage and profit per unit of capital can be written as

$$W_s = (1-\alpha)\left[\alpha\left(\Theta K_s\right)^{\frac{\epsilon-1}{\epsilon}}+1-\alpha\right]^{\frac{1}{\epsilon-1}}, \qquad \pi_{s,j} = \alpha\theta_j\left[\alpha\left(\Theta K_s\right)^{\frac{\epsilon-1}{\epsilon}}+1-\alpha\right]^{\frac{1}{\epsilon-1}}\left(\Theta K_s\right)^{-\frac{1}{\epsilon}} \tag{123}$$

The derivative of the wage is given by

$$\frac{\partial W_s}{\partial K_s} = (1-\alpha)\frac{\alpha\Theta}{\epsilon}\left[\alpha\left(\Theta K_s\right)^{\frac{\epsilon-1}{\epsilon}}+1-\alpha\right]^{\frac{1}{\epsilon-1}-1}\left(\Theta K_s\right)^{-\frac{1}{\epsilon}}, \tag{124}$$

and the derivative of $\pi_{s,j}$ is given by

$$\frac{\partial \pi_{s,j}}{\partial K_s}K_s = -(1-\alpha)\frac{\alpha\theta_j}{\epsilon}\left[\alpha\left(\Theta K_s\right)^{\frac{\epsilon-1}{\epsilon}}+1-\alpha\right]^{\frac{1}{\epsilon-1}-1}\left(\Theta K_s\right)^{-\frac{1}{\epsilon}}. \tag{125}$$

Following similar steps to the case with Cobb-Douglas production function, we find that the derivative of $V(\Delta)$ is given by

$$V'(0) = \beta\mathbb{E}\left[u'(C_s^i)\left(\frac{\partial R_s^{a,i}}{\partial\Delta}K_s + \frac{\partial W_s}{\partial\Delta}\right)\right] \tag{126}$$

$$= \beta\mathbb{E}\left[\frac{1-\tilde{\alpha}_s}{\epsilon}Cov_s(C_s^i), R_s^{a,i})(\kappa_0+\kappa_1\varphi_s^1)\right], \tag{127}$$

where $1-\tilde{\alpha}_s \equiv (1-\alpha)\left[\alpha\left(\Theta K_s\right)^{\frac{\epsilon-1}{\epsilon}}+1-\alpha\right]^{-1}$ is the labor share in state s.

We obtain two differences with respect to the formula in the baseline model. First, the labor share varies across states in the CES case. Second, the welfare impact of the intervention is amplified if the elasticity of substitution ϵ is less than one, and the effect is dampened if $\epsilon > 1$. For instance, Oberfield and Raval (2014) estimates an elasticity of 0.7, which gives an amplification of around 40%.

B.1.3 Endogenous participation choice

We consider next the case in which the participation parameter ϕ is endogenous. Investors can now choose the optimal level of ϕ subject to a cognitive cost. This cost could reflect costs related to acquisition and processing of information or simply a disutility associated with having to pay attention to a larger number of firms. Formally, we introduce a cognitive cost $\mathcal{I}(\phi)\sigma_\theta^2$, where $\mathcal{I}(\cdot)$ is convex and satisfies $\mathcal{I}'(0) = 0$ and $\lim_{\phi\to 1}\mathcal{I}'(\phi) = \infty$. The cognitive cost then increases with the fraction of firms the investors has to pay attention to as well as the amount of uncertainty on each firm, so the cost to learn about the firms vanishes when there is no uncertainty about them.

The investor's problem can now be written in two steps. First, the optimal portfolio choice for a given ϕ. Denote the value function obtained at this stage by $W(\phi)$. Second, a market participation choice, which consists of maximizing $W(\phi) - \mathcal{I}(\phi)\sigma_\theta^2$.

The asymptotic expansion of $W(\phi)$ is given by

$$W(\phi) = W^* - u'(C_0^*)\hat{I}\sigma_\theta^2 + \beta\mathbb{E}\left[u'(C_s^*)\hat{C}_s + \frac{1}{2}u''(C_s^*)\tilde{C}_s^2(1-\phi)^3\right]\sigma_\theta^2 + o(\sigma_\theta^2) \tag{128}$$

where $W^* = u(C_0^*) + \beta\mathbb{E}[u(C_s^*)]$.

Using the optimality conditions for $\hat{I}$ and $\hat{\chi}$, the first-order condition for ϕ can be written as

$$\gamma\beta\mathbb{E}\left[(C_s^*)^{-(\gamma+1)}\right]\frac{3}{2}(1-\phi)^2 = \mathcal{I}'(\phi), \tag{129}$$

where there exists a unique solution $0 < \phi^* < 1$ for the first-order condition above, given the assumptions on $\mathcal{I}(\cdot)$.

Consider next the welfare impact of a given perturbation:

$$V(\Delta) = \max_{\Omega_j^i,\phi}\left\{u\left(E_0 - \sum_{k=0}^{1} I^k(\Delta)\right) + \beta\mathbb{E}\left[u\left(\left(\int_{0-}^{1} R_{s,j}^a(\Delta)d\Omega_j^i + \tau_s(\Delta)\right)K_s(\Delta)\right)\right] - \mathcal{I}(\phi)\sigma_\theta^2\right\},$$

Applying an envelope argument on ϕ, the derivative $V'(0)$ is identical to the one in the case where ϕ is exogenous. Hence, our results apply directly to this case as well. Moreover, the value of ϕ that solves the problem above for $\Delta = 0$ coincides with the one in laissez-faire. Hence, starting from the laissez-faire allocation, the planner has no incentives at the margin to distort the investor's participation decision.

C Appendix for Section 4

C.1 Data description

Variable definitions. We follow Welch (2019) in calculating market betas. Specifically, for each stock-month, we obtain daily return data for the previous 60 months and we winsorize the stock's daily excess return at $(1 \pm 3)\times$ market excess return. Then we run a weighted-least-squares (WLS) univariate regression of this stock's winsorized excess return on the market excess return; the weight is computed according to a decay rate of 2/252 per day (that is, older observations are given lower weights). The WLS slope coefficient is our estimate of market beta (β^W). The average β^W in our sample is 0.8, consistent with Welch (2019).

We compute the market capitalization (ME) for a company by aggregating the market value of all its outstanding shares (which is equal to the product of the price per share and the number of shares outstanding—both variables come from the CRSP data). Then we assign a firm's ME to its stocks. [43] We convert ME into real terms using the CPI index to make it more comparable across time. The median stock in our sample has a ME of around 46 million real dollars.

We follow Fama and French (1992) in calculating the book-to-market (BM) ratio, which is the book value of equity divided by the market value of equity; both variables are calculated using fiscal yearend information from the Compustat database.[44] For each firm, we match the BM ratio for a fiscal year ending in year $t-1$ to its monthly stock returns from July of year t through June of year $t+1$; this is

[43]Note that, for stocks whose issuing firms have multiple share classes, they are assigned the ME of their issuing firms, which are *not* equal to their own market values.

[44]For the market value of equity, if it is not available from an annual accounting record, we calculate it using the subsequent fiscal quarter's information.

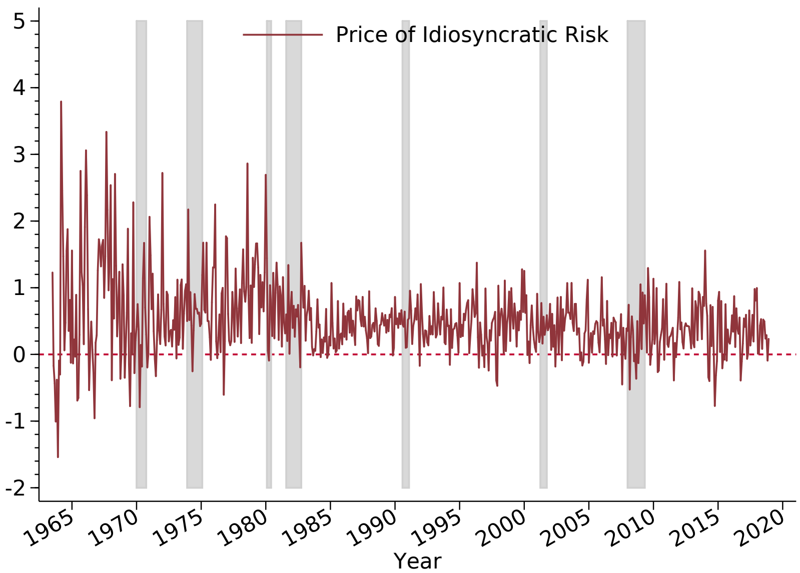


Figure 5: Price of idiosyncratic risk. This figure displays month-by-month estimates of the price of idiosyncratic risk as measured by first-stage Fama-MacBeth slope coefficients for expected idiosyncratic variance (λ_{ivar}). A selection of other characteristics is also included in the Fama-MacBeth regressions to control for standard risks. Shaded areas indicate U.S. recessions identified by NBER.

to ensure that a BM ratio is known before the returns it predicts. In our sample, the median stock has a BM ratio of 0.66.

We measure a stock's past performance by a six-month cumulative gross return. For each month t, we compute, stock by stock, the buy-and-hold compound gross return from month $t-7$ through $t-2$; the adjacent month $t-1$ is excluded to avoid short-term reversals that are likely caused by trading frictions. Holding a median stock in our sample for six months provides a total return of around 3.3%.

Lastly, we compute two measures of liquidity and its variability following Chordia, Subrahmanyam and Anshuman (2001). For each stock-month, we calculate the average of the monthly share turnovers (that is, the share volume divided by the total shares outstanding) over the previous 36 months ($TURN$) as well as the coefficient of variation for share turnovers over that period ($CVTURN$). In our sample, the median stock experiences average monthly turnover of 5.22%, and the corresponding coefficient of variation is 59.32%.

Stability of the price of risk estimation

We plot, in Figure 5, the first-stage Fama-MacBeth slope coefficients for expected idiosyncratic variance (λ_{id}), which can be construed as a measure of idiosyncratic risk premium; all the other characteristics are also included to control for standard risks. As shown, the idiosyncratic risk premium witnessed significant variations in the 1960s and 1970s, but has become fairly stable ever since. There is no discernible cyclical pattern whatsoever. This is consistent with our model in which the idiosyncratic risk premium is equal to the product of γ and ϕ, both of which are constant.

C.2 Proof of lemma 1

Proof. This derivation follows Campbell et al. (2001) closely and it is provided for completeness. Let $r_{j,t}$ denote the return on firm j, $r_{m,t} = \sum_i w_{m,i} r_{i,t}$ the return on the market, where w_m denote the market portfolio weights, and $\beta_{j,t}$ firm j's (conditional) market beta. By definition of market beta, we obtain that $r_{j,t+1} = \beta_{j,t} r_{m,t+1} + \tilde{v}_{j,t+1}$, where $Cov(r_{m,t+1}, \tilde{v}_{j,t+1}) = 0$. Finally, define $v_{j,t+1} \equiv r_{j,t+1} - r_{m,t+1} = (\beta_{j,t} - 1) r_{m,t+1} + \tilde{v}_{j,t+1}$.

The variance of returns can be written as

$$\begin{aligned} Var_t[r_{j,t+1}] &= Var_t[r_{m,t+1}] + Var_t[v_{j,t+1}] + 2Cov_t(r_{m,t}, v_{j,t}) \\ &= Var_t[r_{m,t+1}] + Var_t[v_{j,t+1}] + 2(\beta_{j,t} - 1) Var_t(r_{m,t+1}) \end{aligned} \tag{130}$$

Let $\overline{\sigma}_t^2 \equiv \sum_j w_{m,j} Var_t[r_{j,t+1}]$ and $\overline{\sigma}_{id,t}^2 \equiv \sum_j w_{m,j} Var_t[v_{j,t+1}]$, then

$$\overline{\sigma}_t^2 = \sigma_{m,t}^2 + \overline{\sigma}_{id,t}^2, \tag{131}$$

where $\sigma_{m,t}^2 = Var_t[r_{m,t+1}]$ and we used $\sum_j \omega_{m,j} \beta_{j,t} = 1$. □

C.3 Derivation of the share invested in the risky technology

The optimality condition for the risky technology can be written as

$$\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_{s,j}^a \varphi_s^e\right] = 0 \Rightarrow \mathbb{E}[\varphi_s^e] = -Cov\left(\mathbb{E}_s\left[\beta \frac{(C_s^i)^{-\gamma}}{C_0^{-\gamma}} R_{s,j}^a\right], \varphi_s^e\right). \tag{132}$$

From equation (11) and $\overline{C}_s = \overline{R}_s^a K_s$, we obtain the approximate expression

$$\mathbb{E}[\varphi_s^e] \approx Cov\left(\gamma k_s + (\gamma - 1)\log \overline{R}_s^a - \frac{\gamma(\gamma-1)}{2}\phi_u \sigma_s^2, \varphi_s^e\right). \tag{133}$$

Using $k_s = \log(1 + \chi\varphi_s^e) + \log I \approx \chi\varphi_s^e + \log I$, we obtain

$$\mathbb{E}[\varphi_s^e] = \chi\gamma\sigma_\varphi^2 + (\gamma - 1)Cov\left(\log \overline{R}_s^a - \frac{\gamma}{2}\phi_u\sigma_s^2, \varphi_s^e\right). \tag{134}$$

Rearranging the expression above, we can solve solve for χ

$$\chi = \frac{\mathbb{E}[\varphi_s^e]}{\gamma\sigma_\varphi^2} - \left(1 - \frac{1}{\gamma}\right)\left[\frac{Cov(\log \overline{R}_s^a, \varphi_s^e)}{\sigma_\varphi^2} - \frac{\gamma\phi_u}{2}\frac{Cov(\sigma_s^2, \varphi_s^e)}{\sigma_\varphi^2}\right]. \tag{135}$$

D Appendix for Section 5

D.1 Appendix for Subsection 5.1

In this appendix, we provide the remaining elements necessary for a description of the economy subject to financial regulation, its equilibrium, and the proof of Proposition 8, which establishes the condition for implementation of an allocation with financial regulation.

The modified investor's problem – The investor's problem in the regulated economy is

$$\max_{C_0^i, \{\Omega_j^i\}_{j\in[0,1)}} u\left(C_0^i\right) + \beta\mathbb{E}\left[u\left(C_s^i\right)\right], \tag{136}$$

subject to a non-negativity condition on consumption, the participation constraint $\int_{\mathcal{P}^i} d\Omega_j^i = 1$, and budget constraint

$$C_s^i = R_s^i(E_0 - T - C_0^i) + T_s^w,$$

where

$$R_s^i \equiv \Psi^i \int_{0-}^1 \frac{R_{s,j}^a K_{s,j} - D_j}{P_{e,j}} d\Omega_j^i + \left(1 - \Psi^i\right) \frac{1}{P_d},$$

is the return on the investor's portfolio, Ψ^i is the portfolio weight on risky assets, Ω_j^i is the equity portfolio distribution, T is a lump-sum levy used to finance the debt tax shield, and T_s^w is a lump-sum transfer from workers.

The modified equilibrium definition – An *allocation* is given by consumption and portfolio decisions for investors, $\left(C_0^i, \Psi^i, [\Omega_j^i, \Omega_{d,j}^i]_{j\in[0,1)}\right)$ for $i \in [0,1)$, investment and labor demand decisions for firms, $\left(I_j^0, I_j^1, L_{l,j}, L_{h,j}\right)$ for $j \in [0,1)$, and workers' consumption, (C_l^w, C_h^w). A *competitive equilibrium* is defined as an allocation, asset prices $\left(P_{e,j}, P_{d,j}\right)$ for each firm j, and wages W_s for each state s such that:

i. Consumption and portfolio decisions, $\left(C_0^i, \Psi^i, \{\Omega_j^i\}_{j\in[0,1)}\right)$, solve Problem (136) for each $i \in [0,1)$.

ii. Investment and debt issuance decisions solve Problem (18) given $M_{s,j} = \frac{1}{\phi}\int_{\{i:j\in\mathcal{P}^i\}} \beta \frac{u'(C_s^i)}{u'(C_0^i)} di$, and labor demand is given by 2.

iii. Asset markets for equity and debt clear.

iv. The goverment's budget at $t = 0$ is balanced, with $T = \tau^d D$.

v. Worker consumption in each state $s \in \mathcal{S}$ is given by $C_s^w = W_s - T_s^w$.

vi. The labor market clears at each $s \in \mathcal{S}$, i.e. $\int_0^1 L_{s,j} dj = 1$.

vii. Consumption goods markets clear, i.e., $C_0 + \sum_{k=0}^1 I^k = E_0$,where $I^k \equiv \int_0^1 I_j^k dj$ for $k = \{0,1\}$, and at each $s \in \mathcal{S}$

$$C_s^w + \int_0^1 C_s^i di = \int_0^1 (\theta_j K_{s,j})^\alpha L_{s,j}^{1-\alpha} dj + (1-\delta) K_s,$$

where $K_{s,j} = \sum_{k=0}^1 \varphi_s^k I_j^k$ and $K_s = \int_0^1 K_{s,j} dj$.

Definition 1. *An allocation features an implicit investment subsidy whenever, for each $j \in [0,1)$,*

$$1 \geq \mathbb{E}\left[M_{s,j}\left(R^a_{s,j}\left(1+\chi\varphi^e_s\right)\right)\right].$$

An allocation features an implicit risk-taking tax whenever, for each $j \in [0,1)$

$$\mathbb{E}\left[M_{s,j}R^a_{s,j}\varphi^1_s\right] \geq \mathbb{E}\left[M_{s,j}R^a_{s,j}\varphi^0_s\right].$$

Proposition 8 (Implementation). *A symmetric allocation $\left(C_0, \left\{I^k\right\}_{k=0,1}, \left\{\Omega^0_j\right\}_{j\in[0,1)}, C^i_s, C^w_s\right)$ with an implicit investment subsidy and an implicit risk-taking tax can be implemented with a set of subsidies $\left\{\tau^k_s, \tau^d\right\}$ and financial regulation with risk weights $\left\{\omega^k\right\}_{k=0,1}$ whenever it satisfies:*

i. Feasibility with $E_0 = \sum_k I^k + C_0$ and $K_s = \sum_k \varphi^k_s I^k$.

ii. The distribution of $t=1$ consumption is given by

$$C_{i,s} = \int_{0-}^{1} \frac{R^a_{s,j}K_{s,j}}{\sum_k I^k} d\Omega^i_{e,j} K_s + T^w_s,$$

where $\Omega^i_{e,j}$ ensures that for every $\left(j, j^{'}\right) \in \mathcal{P}^i$

$$E\left[u^{'}\left(C_{i,s}\right)R^a_{s,j}K_s\right] = E\left[u^{'}\left(C_{i,s}\right)R^a_{s,j^{'}}K_s\right]$$

and

$$C^w_s = (1-\alpha)\,\Theta^\alpha K^\alpha_s - T^w_s,$$

for some T^w_s.

Furthermore, $\tau^d > 0$ and $\omega_1 > \omega_0$.

Proof of Proposition 8. Take an allocation that satisfies the requirements of the proposition. Define $I = \sum_k I^k$ and $\chi = I^1/I$. Let $d = \frac{D}{I}$ and take any $0 < d \leq (1-\delta)\left(1-\chi\left(1-\varphi^1_L\right)\right)$. We verify that we can find a system of subsidies and risk weights that satisfies all the conditions for an equilibrium.

Investor optimality. From the investor's side, we obtain, for savings,

$$1 = \beta E\left[\frac{u^{'}\left(C^i_s\right)}{u\left(C_0\right)}R^i_s\right],$$

for the portfolio shares

$$\beta E\left[\frac{u^{'}\left(C^i_s\right)}{u\left(C_0\right)}R_e\right] = \frac{1}{P_d}\beta E\left[\frac{u^{'}\left(C^i_s\right)}{u\left(C_0\right)}\right],$$

where R_e is the optimal equity portfolio's (random) return. Together, these are equivalent to

$$P_e = \beta E\left[\frac{u^{'}\left(C^i_s\right)}{u\left(C_0\right)}\left(R^a_{s,j}\left(1+\chi\varphi^e_s\right)-d\right)\right]I, \tag{137}$$

for each $j \in \mathcal{P}^i$ and

$$P_d = \beta E\left[\frac{u^{'}\left(C_s^i\right)}{u\left(C_0\right)}\right]. \tag{138}$$

Firm optimality. As discussed in Section 5, investment and capital structure decisions are made to maximize the joint surplus of the intermediary-firm relationship. We seek to construct an allocation in which the debt constraint in (19) is not binding in the firm's problem. In such a situation, the problem can be rewritten as

$$\max_{d,I,\chi\geq 0}\left\{P_d\left(1+\tau^d\right)d-1+\mathbb{E}\left[M_{s,j}\left(R_{s,j}^a\left(1+\chi\varphi_s^e\right)-d\right)\right]\right\}I,$$

s.t.

$$1-P_d d\geq \omega^0\left(1-\chi\right)+\omega^1\chi.$$

Its first-order conditions give us, for I, χ and d, respectively,

$$P_d\left(1+\tau^d\right)d+\mathbb{E}\left[M_{s,j}\left(R_{s,j}^a\left(1+\chi\varphi_s^e\right)-d\right)\right]=1, \tag{139}$$

$$\frac{\mathbb{E}\left[M_{s,j}R_{s,j}^a\varphi_s^e\right]}{\omega^1-\omega^0}=\frac{\mu^{rw}}{I}\geq 0, \tag{140}$$

and

$$P_d\left(1+\tau^d\right)-\mathbb{E}\left[M_{s,j}\right]=\frac{\mu^{rw}}{I}\geq 0. \tag{141}$$

Additionally, it is required that

$$1-P_d d=\omega^0\left(1-\chi\right)+\omega^1\chi, \tag{142}$$

and

$$\left(1-\delta\right)\left(1-\chi\left(1-\varphi_L^1\right)\right)\geq d. \tag{143}$$

Labor market equilibrium and worker consumption. Similarly to laissez-faire, optimality and labor market clearing can be ensured under $W_s=\left(1-\alpha\right)\Theta^\alpha K_s^\alpha$. In the presence of the lump-sum tax, we have

$$C_s^w=\left(1-\alpha\right)\Theta^\alpha K_s^\alpha-T_s^w. \tag{144}$$

Market-clearing for equity and debt. Market clearing for equity requires that, for aggregates,

$$\Psi\left(E_0-T-C_0\right)=P_e, \tag{145}$$

and

$$\left(1-\Psi\right)\left(E_0-T-C_0\right)=IP_d d. \tag{146}$$

Verification of implementability. We seek to find $P_e, P_d, \tau^d, \left\{\omega^k\right\}$ that support the candidate allocation and $d>0$ as an equilibrium. Notice first that, from (137) and (138), asset prices are given as a function of the allocation. Equation (138) together with Equation (139) and the fact that $E\left[M_{s,j}\right]=P_d$

delivers

$$\tau^d d + \mathbb{E}\left[M_{s,j}\left(R^a_{s,j}\left(1+\chi\varphi^e_s\right)\right)\right] = 1, \tag{147}$$

which pins down τ^d. Notice that, because the allocation features an implicit investment subsidy, $\tau^d > 0$. Equation (141) implies that $\frac{\mu^{rw}}{I} = \tau^d P_d \geq 0$. Then, we can use Equation (140) to establish that

$$\omega^1 - \omega^0 = \frac{\mathbb{E}\left[M_{s,j}R^a_{s,j}\varphi^e_s\right]}{\tau^d P_d} \geq 0.$$

Lastly, we can obtain ω_0 from Equation (142). Set $T = I\tau^d d$ and use Equation (145) to solve for Ψ. It then follows that, adding (145) and (146), we obtain

$$(E_0 - T - C_0) = P_e + IP_d d.$$

Using feasibility at date $t = 0$ and (147), we verify that this equation holds, proving equality in Equation 146. □

D.2 Appendix for Subsection 5.2

Proof of Proposition 6. We write the relaxed planning program as

$$\max_{I,\chi,\Omega^i_j,\{T^w_s\}_s} u\left(E_0 - I\right) + \beta\mathbb{E}\left[u\left(R^{a,i}_s K_s + T^w_s\right)\right], \tag{148}$$

subject to the limited participation constraint (4) and

$$E\left[u^w\left((1-\alpha)\left(\Theta K_s\right)^\alpha - T^w_s\right)\right] \geq \underline{u}^w, \tag{149}$$

where $R^a_{j,s} = 1 - \delta + \alpha\theta_j\left(\Theta K_s\right)^{\alpha-1}$, $R^{a,i}_s = \int_{\mathcal{P}^i} R^a_{s,j} d\Omega^i_j$, and $K_s = (1+\chi\varphi^e_s)I$.

The first-order condition with respect to T^w_s is

$$\beta\mathbb{E}_s\left[u'(C^i_s)\right] = \mu^w u^{w\prime}(C^w_s). \tag{150}$$

The optimality condition for the portfolio allocation is

$$\mathbb{E}\left[u'\left(C^i_s\right)R^a_{s,j}K_s\right] = \mathbb{E}\left[u'\left(C^i_s\right)R^{a,i}_s K_s\right]. \tag{151}$$

The first-order condition for I is

$$-u'(C_0) + \beta\mathbb{E}\left[u'(C^i_s)R^{a,i}_s\frac{K_s}{I}\right] + \beta\mathbb{E}\left[u'(C^i_s)\frac{\partial R^{a,i}_s}{\partial I}K_s\right] + \mu^w\mathbb{E}\left[u^{w\prime}(C^w_s)\alpha(1-\alpha)\Theta^\alpha K^{\alpha-1}_s\frac{K_s}{I}\right] = 0, \tag{152}$$

where

$$\frac{\partial R^{a,i}_s}{\partial I} = -(1-\alpha)\alpha\theta^i\Theta^{\alpha-1}K^{\alpha-2}_s\frac{K_s}{I}, \qquad \theta^i \equiv \int_{\mathcal{P}^i}\theta_j d\Omega^i_j. \tag{153}$$

Using the optimality condition for T_s^w, then we obtain

$$1 = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_s^{a,i} \frac{K_s}{I}\right] - (1-\alpha)\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} \left(R_s^{a,i} - \overline{R}_s\right) \frac{K_s}{I}\right]. \tag{154}$$

The second term in the above expression is the risk externality for investment

$$IRE_I \equiv -(1-\alpha)\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} \left(R_s^{a,i} - \overline{R}_s\right) \frac{K_s}{I}\right]. \tag{155}$$

We can rewrite the above expression as

$$1 = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_s^{a,i} \frac{K_s}{I(1 - IRE_I)}\right]. \tag{156}$$

From the optimality condition for the portfolio allocation, we can replace $R_s^{a,i}$ by $R_{s,j}^a$, $j \in \mathcal{P}^i$,

$$1 = \mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_{s,j}^{a} \frac{K_s}{I(1 - IRE_I)}\right]. \tag{157}$$

The optimality condition for the share of the risky technology is

$$\beta\mathbb{E}\left[u'(C_s^i) R_s^{a,i} \varphi_s^e\right] + \mu^w \mathbb{E}\left[u'^{w'}(C_s^w)\alpha(1-\alpha)\Theta^\alpha K_s^{\alpha-1}\varphi_s^e\right] + \beta\mathbb{E}\left[u'(C_s^i)\frac{\partial R_s^{a,i}}{\partial \chi}\frac{K_s}{I}\right] = 0, \tag{158}$$

where

$$\frac{\partial R_s^{a,i}}{\partial \chi} = -(1-\alpha)\alpha\theta^i\Theta^{\alpha-1}K_s^{\alpha-2}\varphi_s^e I. \tag{159}$$

Using the optimality condition for T_s^w, we obtain

$$\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} R_s^{a,i} \varphi_s^e\right] = IRE_\chi, \tag{160}$$

where

$$IRE_\chi \equiv (1-\alpha)\mathbb{E}\left[\beta \frac{u'(C_s^i)}{u'(C_0)} \left(R_s^{a,i} - \overline{R}_s^a\right) \varphi_s^e\right]. \tag{161}$$

□