

Under-diversification and idiosyncratic risk externalities ^{*}

Felipe S. Iachan

Dejanir Silva

Chao Zi [†]

May, 2021

Abstract

We study the effects of idiosyncratic uncertainty on asset prices, investment, and welfare. We consider an economy with two main components: under-diversification and endogenous, countercyclical idiosyncratic risk. 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. We assess their magnitude empirically. The optimal allocation can be implemented through financial regulation using a tax benefit on debt and risk-weighted capital requirements.

KEYWORDS: Idiosyncratic risk, Under-diversification, Risk-taking, Pecuniary externalities

JEL CLASSIFICATION: E22, G12, G18.

^{*}Iachan: FGV EPGE. E-mail: felipe.iachan@fgv.br. Silva: University of Illinois at Urbana-Champaign. E-mail: dejanir@illinois.edu. Corresponding author. Mailing address: 1206 South Sixth Street, 461 Wohlers Hall, Champaign, IL, 61820. Phone: 217-300-5782. Fax: 217-244-3102. Zi: Shanghai Advanced Institute of Finance (SAIF), Shanghai Jiao Tong University. E-mail: czi@saif.sjtu.edu.cn.

[†]We would like to thank Antonio Antunes, Victor Duarte, Bernard Herskovic, Tim Johnson, Mahyar Kargar, Dana Kiku, and seminar/conference participants at the Lubramacro, Lubrafin, UIUC, IDB, Fed Board, FSWF, SED, and SBE meetings. We are particularly grateful to the editor, William Schwert, and two anonymous referees whose invaluable comments significantly contributed to the paper. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

1. Introduction

There is significant under-diversification of idiosyncratic risk.¹ Frictions in diversification enable idiosyncratic risk to affect asset prices, distort corporate investment and funding decisions, and even create economy-wide fluctuations.² Nonetheless, little is known about whether and how policymakers could alleviate the inefficiencies created by idiosyncratic risk or respond to its cyclical properties.

In this paper, we study the effects of idiosyncratic uncertainty on asset prices, investment, and welfare. In particular, we consider the inefficiency of the equilibrium allocation and its policy implications. We analyze this question in the context of a production asset-pricing model with two main components: (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 endogenous cyclical properties of risk play an important role, as they are shaped by investment decisions and influenced by regulation.

Our main result is that the economy is subject to a new form of pecuniary externality, which we call 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 underinvestment and excessive aggregate risk-taking in a laissez-faire economy. Moreover, we derive and apply sufficient statistics that resort to asset pricing data to quantify the importance of these two forms of investment inefficiency. Finally, we show how financial regulation can be used to address these inefficiencies.

We consider a two-period model with firms, investors, and workers. Firms allocate investment across a riskless and a risky technology. The payoff of the risky technology is the only source of aggregate risk. In the second period, firms combine capital with labor using a neoclassical production function subject to idiosyncratic productivity shocks. Capital cannot be reallocated once

¹Under-diversification is pervasive for entrepreneurs and outside investors. For example, [Himmelberg et al. \(2000\)](#) document that entrepreneurs hold a large fraction of wealth in their own companies. Under-diversification in investors' portfolios has been documented by [Blume and Friend \(1975\)](#), [Kelly \(1995\)](#), [Polkovnichenko \(2005\)](#), and [Calvet et al. \(2007\)](#).

²See, for example, [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 ([Bloom, 2009](#); [Gabaix, 2011](#); [Acemoglu et al., 2012](#); [Christiano et al., 2014](#)).

it is installed. Investors and firms are located on a circle. Firms are subject to local productivity shocks, but average productivity across all locations is constant.

In the first period, investors choose how much to consume and select an equity portfolio subject to a limited-participation constraint. Investors have access only to firms located in a given neighborhood. This friction can be interpreted as capturing the fact that investors' portfolios are concentrated 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 size of the neighborhood to which investors have access acts as a measure of diversification. For example, if an investor has access to the whole circle, she would be able to perfectly diversify idiosyncratic risk. At the other extreme, if an investor can invest only in a single firm, she would fully bear the idiosyncratic risk of that firm, as, for example, in the entrepreneurship model of [Chen et al. \(2010\)](#).

In the second period, workers inelastically supply labor and consume the final good. 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) the dispersion in the volume produced, as 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 in [Donangelo et al. \(2019\)](#).

This mechanism has important asset pricing implications. First, the model generates the synchronized idiosyncratic volatility observed in [Herskovic et al. \(2016\)](#), even without an assumption of state-dependent productivity dispersion. Second, the model generates a negative premium for exposure to states where idiosyncratic volatility is high, consistent again with the evidence 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 in the level of idiosyncratic volatility. 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, the model generates a positive idiosyncratic variance risk premium, that is,

a positive difference between expected idiosyncratic variance under the risk-neutral and the physical probabilities. Multiple studies document a positive aggregate variance risk premium and, in Section 4, we show evidence of an idiosyncratic variance risk premium.³ The expected return on the firm's equity can be decomposed into an aggregate risk premium, which is proportional to the covariance of returns with aggregate consumption, and an idiosyncratic risk premium, which is proportional to the level of idiosyncratic volatility. As in the original model of Merton (1987), we find that, in equilibrium, idiosyncratic risk commands a positive premium. As the price of idiosyncratic risk depends on the degree of under-diversification, we can estimate that degree, in a necessary step to quantitatively assess the welfare implications of idiosyncratic risk.

The investment decisions of firms transmit these asset-pricing effects of under-diversification to the real economy. Idiosyncratic risk leads to a reduction in aggregate risk-taking compared to 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 increases with consumption dispersion. Additionally, this endogenous reduction in aggregate risk-taking counterbalances the presence of idiosyncratic risk, resulting in an *laissez-faire* investment level that can be either above or below the corresponding level in the first-best economy.

We next consider the policy implications of the inefficiencies created by idiosyncratic risk. We study the problem of a social planner that cannot directly control the degree of diversification in private portfolios, but can 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 are hard to be addressed directly by policy.⁴

Our main result is that, in the absence of interventions, the economy is constrained-inefficient. In other words, even a planner who is constrained not to directly increase diversification can bring about welfare improvements. The inefficiency results from a pecuniary externality in investment decisions. First, 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

³See, for example, 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 related literature.

⁴Van Nieuwerburgh and Veldkamp (2010) show 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.

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. This occurs despite a private risk-taking level under laissez-faire that is below the first-best. The direction of the intervention is therefore dictated by the externality, not by a comparison with the first-best.

We employ a sufficient-statistic approach to quantify investment inefficiencies. We derive measures of the welfare implications of alternative investment policies. These measures can be computed directly from asset price data, circumventing the need to estimate the model's complete structure. We first derive an estimate of the welfare gains from an investment increase, starting from a laissez-faire equilibrium. These welfare gains depend on the idiosyncratic risk premium and the idiosyncratic variance risk premium. Similarly, we measure the impact of a reduction in aggregate risk-taking. We show that the gains from this reduction depend on the idiosyncratic variance risk premium and the risk-neutral probabilities.

By assessing these sufficient statistics, we show that there are significant welfare gains from correcting risk externalities. We find 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 from 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 4% 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 optimal design of financial regulation. We introduce a financial intermediary and show that a tax benefit 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 benefit and the optimal risk weights can be related directly to asset prices.

Our paper is related to the classical work on under-diversification including [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. \(2020\)](#), [Dou \(2016\)](#), [Di Tella \(2017\)](#), [Silva and Townsend \(2019\)](#) and [Khorrami \(2019\)](#). Other papers have focused on the corporate finance implications of idiosyncratic risk, including [Chen et al. \(2010\)](#), [Wang et al. \(2012\)](#), and [Chen and Strebulaev \(2018\)](#).

In its normative approach, our work follows a tradition that studies constrained-inefficient allocations and pecuniary externalities stemming from incomplete markets.⁵ We offer the first analysis of how endogenous and cyclical idiosyncratic return risk can lead to constrained-inefficient outcomes and describe the role of policy in mitigating such inefficiencies.

Two ingredients are necessary for the identification of idiosyncratic risk externalities: endogenous idiosyncratic risk, resulting from variable input prices, and uninsurable idiosyncratic shocks with risk-averse investors. Their interaction has not been previously studied. For instance, some studies feature uninsurable idiosyncratic risk, but no feedback from input prices to return risk. For example, the model in [Gromb and Vayanos \(2002\)](#) does not feature production, while [Di Tella \(2019\)](#) does not feature other inputs besides capital. On the polar opposite, other work focuses on models with variable inputs, but no idiosyncratic return risk (e.g., [Itskhoki and Moll, 2019](#); [Bocola and Lorenzoni, 2020](#)). Finally, there are a number of studies on pecuniary externalities with risk-neutral agents (e.g., [Lorenzoni, 2008](#); [Caballero and Lorenzoni, 2014](#); [He and Kondor, 2016](#); [Lanteri and Rampini, 2020](#)). By simultaneously considering endogenous return risk and under-diversified investors, we identify this new source of inefficiencies.

Last, our work is also related to the literature on uninsurable income risk, which includes [Constantinides and Duffie \(1996\)](#), [Brav et al. \(2002\)](#), and [Constantinides and Ghosh \(2017\)](#). Like these papers, we consider a SDF that depends on countercyclical consumption risk, but we focus on rate of return risk in a production economy instead of labor income risk.

⁵See e.g. [Stiglitz \(1982\)](#), [Geanakoplos and Polemarchakis \(1986\)](#), and [Greenwald and Stiglitz \(1986\)](#). For a recent taxonomy of such pecuniary externalities, see [Dávila and Korinek \(2018\)](#).

2. A model of under-diversification and investment allocation

In this section, we examine the implications of the under-diversification of idiosyncratic risk for asset pricing and investment decisions. First, we discuss the environment, then we characterize the equilibrium.

2.1. Environment

We study a finite-horizon economy with two dates, $t = 0, 1$. The economy is populated by workers and investors, 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 number N of ex ante identical agents, who make savings and portfolio decisions at date $t = 0$. Investors are indexed by $i \in \mathcal{I} = \{\frac{1}{N}, \frac{2}{N}, \dots, 1\}$, which represents their location in the circle.

There is also a number N of ex ante identical firms, indexed by $j \in \mathcal{J} = \{\frac{1}{N}, \frac{2}{N}, \dots, 1\}$, their location in the circle. Firms raise equity to finance investment at date $t = 0$. They pay dividends in period one from the proceeds of the production of final goods.

Uncertainty has both an aggregate and an idiosyncratic component. They are both public information once realized. 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 $h(l)$ as the high(low) state, in which production will be endogenously higher(lower).

Firm j also learns its idiosyncratic productivity parameter $\theta_{j,s} \in \mathbb{R}_+$. The productivity shocks $\theta_{j,s}$ are identically distributed across firms, with expected value $\mathbb{E}[\theta_{j,s}] = \Theta$ and variance $\text{Var}[\theta_{j,s}] = \sigma_\theta^2$. Their distribution is independent of the aggregate state s and they are purely idiosyncratic, as they do not affect the economy's aggregate productivity. Formally, the productivity shocks $\theta_{j,s}$ satisfies the following condition:

$$\frac{1}{N} \sum_{j \in \mathcal{J}} \theta_{s,j} = \Theta, \quad (1)$$

almost surely, for each $s \in \mathcal{S}$, which requires the correlation across any pair of firms to be $\rho = -\frac{1}{N-1}$.⁶ An explicit construction of the distribution of $\{\theta_{s,j}\}_{j \in \mathcal{J}}$ based on a discretization of a gamma bridge is provided in Appendix A.1, analogous to the circular economies of [Gârleanu](#)

⁶The variance of the average productivity is $\text{Var}[\frac{1}{N} \sum_j \theta_{s,j}] = (\frac{1}{N} + \frac{N-1}{N} \rho) \sigma_\theta^2$, which is equal to zero when $\rho = -\frac{1}{N-1}$.

et al. (2015) and Gârleanu and Panageas (2018).

2.1.1. 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. 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 (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 much of the recent macro-finance literature (e.g., Brunnermeier and Sannikov, 2014; 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.⁷

2.1.2. Investment allocation problem

The profit maximization by firms can be solved in two stages: a capital-budgeting stage at $t = 0$ and, once uncertainty regarding productivity, effective capital available, and aggregate conditions is resolved, a static profit maximization problem at $t = 1$.

On date $t = 0$, firms choose how much to invest in each 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, $K_{j,s} = \sum_{k=0}^1 \varphi_s^k I_j^k$. The return on assets (ROA) in period 1 is given by $R_{j,s}^a := 1 - \delta + \pi_{j,s}$, where δ is the depreciation rate and $\pi_{j,s}$ is the profit per unit of capital generated by the firm, a function of the idiosyncratic productivity $\theta_{j,s}$ and the aggregate state of the economy s . Let $M_{j,s}$ denote the (equally weighted) average SDF of the firm's shareholders. Given the symmetry of the economy, all shareholders price the payoff of the firm's investment in the same way, so alternative weights when computing

⁷For an analysis of capital budgeting and risk-taking distortions in the presence of financial frictions, see Iachan (2020).

$M_{j,s}$ generate identical results. The problem of the firm can then be written as:

$$\max_{I_j^0, I_j^1 \geq 0} \left\{ - \sum_{k=0}^1 I_j^k + \mathbb{E} \left[M_{j,s} R_{j,s}^a \sum_{k=0}^1 \varphi_s^k I_j^k \right] \right\}. \quad (2)$$

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

$$\mathbb{E} \left[M_{j,s} R_{j,s}^a \varphi_s^k \right] = 1, \quad (3)$$

for $k = 0, 1$.

2.1.3. Profit maximization in period one

Capital cannot be reallocated across firms in period one, so firm j operates $K_{j,s}$ units of capital, regardless of its productivity level. A firm with productivity $\theta_{j,s}$ and $K_{j,s}$ units of capital hires $L_{j,s}$ workers at wage W_s and produces final goods according to the Cobb-Douglas production function $Y_{j,s} = (\theta_{j,s} K_{j,s})^\alpha L_{j,s}^{1-\alpha}$. Each firm chooses how much labor to hire to maximize profits:

$$\max_{L_{j,s}} (\theta_{j,s} K_{j,s})^\alpha L_{j,s}^{1-\alpha} - W_s L_{j,s}. \quad (4)$$

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

$$L_{j,s} = \left[\frac{1-\alpha}{W_s} \right]^{\frac{1}{\alpha}} \theta_{j,s} K_{j,s}, \quad (5)$$

which implies 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 $\pi_{j,s} K_{j,s}$, where the profit per unit of capital is given by:

$$\pi_{j,s} = \alpha \theta_{j,s} \left[\frac{1-\alpha}{W_s} \right]^{\frac{1-\alpha}{\alpha}}. \quad (6)$$

Two aspects of Eq. (6) are worth mentioning. First, profitability is heterogeneous across firms. In a frictionless environment, $\pi_{j,s}$ 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. Sec-

ond, the level and dispersion of firms' profitability are endogenous, as the conditional variance of $\pi_{j,s}$ is:

$$\text{Var}_s [\pi_{j,s}] = \alpha^2 \left[\frac{1 - \alpha}{W_s} \right]^{\frac{2(1-\alpha)}{\alpha}} \sigma_\theta^2. \quad (7)$$

For instance, a drop in wages reduces variable costs, amplifying the effects of changes in productivity, an effect analogous to the one generated by operating leverage.⁸

2.1.4. Under-diversification and the participation constraint

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 $\phi/2$ of her location, so investors have access to firms on an arc of length ϕ , as indicated in Panel A of Fig. 1. The participation set of investor i is then given by $\mathcal{P}_i = \{j \in \mathcal{J} : d(i, j) \leq 0.5\phi\}$, where $d(i, j)$ denotes the distance in the circle between investor i and firm j and the parameter $\phi \in [0, 1]$ controls the degree of under-diversification in the 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. Panel B of Fig. 1 shows the productivity distribution of an individual firm and the distribution of average productivity of firms in the investor's portfolio in the case where $0 < \phi < 1$. As we increase the fraction of firms in the investor's portfolio, the dispersion in average productivity is reduced, so ϕ provides a measure of the degree of diversification obtained by investors.

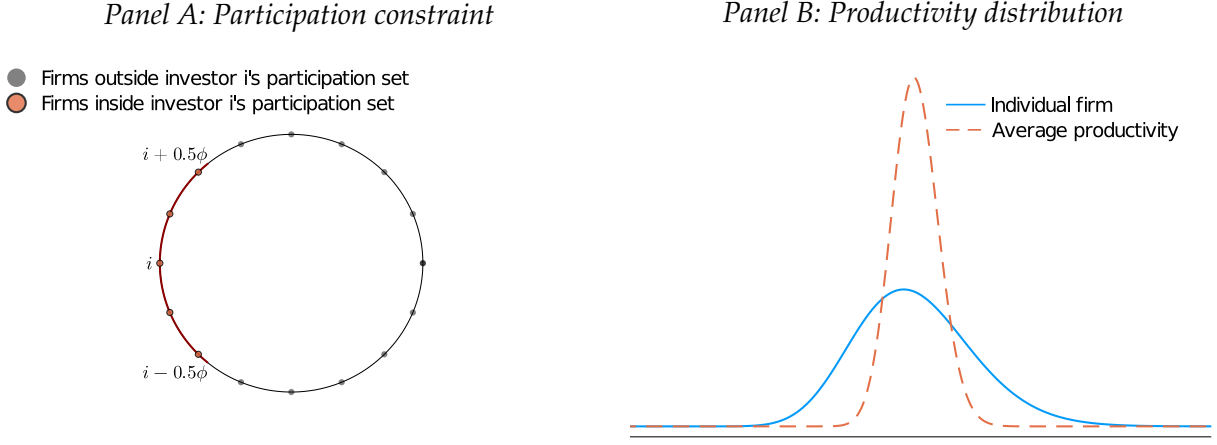
2.1.5. Investor's consumption and savings

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. Let $\omega_{i,j}$ denote the investor's portfolio weight on firm j .

Investors have CRRA preferences with coefficient of risk aversion γ . The investor's problem

⁸Formally, variable costs and production are proportional to productivity, $VC_{j,s} = \overline{VC}_s \theta_j$ and $Y_{j,s} = \overline{Y}_s \theta_j$, so the dispersion in profits is $\sigma_{\pi,s} = (\overline{Y}_s - \overline{VC}_s) \sigma_\theta$. Lower wages increase the margin $\overline{Y}_s - \overline{VC}_s$, amplifying the effect of σ_θ .

Fig. 1. Participation constraint and productivity distribution.



is:

$$\max_{C_{i,0}, \{\omega_{i,j}\}} u(C_{i,0}) + \beta \mathbb{E} [u(C_{i,s})], \quad (8)$$

subject to a non-negativity condition on consumption, limited participation $\omega_{i,j} = 0$ for $j \notin \mathcal{P}_i$, and

$$C_{i,s} = R_{i,s}^p (E_0 - C_{i,0}), \quad R_{i,s}^p := \sum_{j \in \mathcal{P}_i} \omega_{i,j} \frac{R_{j,s}^a K_{j,s}}{P_j}, \quad (9)$$

given $\sum_{j \in \mathcal{J}} \omega_{i,j} = 1$, where $R_{i,s}^p$ is the return on investor i 's portfolio and P_j is firm j 's share price.

The optimality conditions for initial consumption and portfolio weights are given by:

$$1 = \mathbb{E} \left[M_{i,s} R_{i,s}^p \right], \quad (10)$$

and

$$P_j = \mathbb{E} \left[M_{i,s} R_{j,s}^a K_{j,s} \right], \quad (11)$$

for all $j \in \mathcal{P}_i$, where $M_{i,s}$ denotes SDF for investor i and it is given by:

$$M_{i,s} = \beta \left(\frac{C_{i,s}}{C_{i,0}} \right)^{-\gamma}, \quad (12)$$

using $u(C) = \frac{C^{1-\gamma}-1}{1-\gamma}$.

2.1.6. Workers and equilibrium definition

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

$$C_{w,s} = W_s. \quad (13)$$

Let's now define the equilibrium. An allocation is given by consumption and portfolio decisions for investors, $(C_{i,0}, \{\omega_{i,j}\}_{j=1}^N)$ for $i \in \mathcal{I}$, investment and labor demand decisions for firms, $(I_j^0, I_j^1, L_{l,j}, L_{h,j})$ for $j \in \mathcal{J}$, and workers' consumption, $(C_{w,l}, C_{w,h})$. A competitive equilibrium consists of an allocation, asset prices P_j for each firm j , and wages W_s for each state s , such that investors and firms optimize, workers consume their income, and there is market clearing for the stocks of firm $j \in \mathcal{J}$, for consumption goods in period 0,

$$\frac{1}{P_j} \sum_{i \in \mathcal{I}} (E_0 - C_{i,0}) \omega_{i,j} = 1, \quad \sum_{i \in \mathcal{I}} C_{i,0} + \sum_{j \in \mathcal{J}} (I_j^0 + I_j^1) = NE_0, \quad (14)$$

and for labor and consumption goods in state $s \in \mathcal{S}$,

$$\sum_{j \in \mathcal{J}} L_{j,s} = 1, \quad C_{w,s} + \sum_{i \in \mathcal{I}} C_{i,s} = \sum_{j \in \mathcal{J}} (Y_{s,j} + (1 - \delta)K_{s,j}). \quad (15)$$

2.2. Equilibrium characterization

We next consider the characterization of the equilibrium. We focus on a symmetric equilibrium, in which $C_{i,0} = C_0$, $I_j^k = I^k$, $P_j = P$, and $K_{j,s} = \frac{K_s}{N}$ at each $s \in \mathcal{S}$.

2.2.1. Aggregate production, wages, and returns

Taking the ratio of the labor demand for firm j , given in Eq. (5), and the aggregate labor demand across all firms, we obtain $L_{j,s} = \frac{\theta_{j,s}}{N\Theta}$. We can then compute aggregate output as:

$$Y_s = \sum_{j \in \mathcal{J}} (\theta_j K_{s,j})^\alpha L_{j,s}^{1-\alpha} = (\Theta K_s)^\alpha. \quad (16)$$

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

$$W_s = (1 - \alpha) (\Theta K_s)^\alpha. \quad (17)$$

Plugging the wage into Eq. (6), we obtain the return on assets:

$$R_{j,s}^a = 1 - \delta + \alpha \theta_j (\Theta K_s)^{\alpha-1}, \quad (18)$$

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

In equilibrium, the stock price satisfies $P = \sum_{k=0}^1 I^k$ and, as a consequence, equals the replacement cost of capital. Therefore, Tobin's q is one. The fact that q is equal to one allows us to distinguish the inefficiencies we find from those based on the interaction of the price of capital with financial constraints, as in, for example, He and Kondor (2016) and Jeanne and Korinek (2019). The return on investing in firm j is given by the product of the return on assets and the return on investment, as:

$$R_{j,s} = R_{j,s}^a \bar{\varphi}_s, \quad (19)$$

where $\bar{\varphi}_s := \frac{\sum_{k=0}^1 \varphi_s^k I^k}{\sum_{k=0}^1 I^k}$ denotes the average return on the investment technologies in state s .

2.2.2. Equilibrium portfolio and under-diversification

In a symmetric equilibrium, the returns on all pairs of firms share the same joint distributions. As a consequence, the investor's portfolio problem is solved by an equally-weighted equity portfolio within her participation set. Formally, $\omega_{i,j} = \frac{1}{|\mathcal{P}_i|}$ for each $j \in \mathcal{P}_i$ and zero otherwise, where $|\mathcal{P}_i|$ denotes the number of firms in \mathcal{P}_i . The return on this equally-weighted equity portfolio is given by:

$$R_{i,s}^p = \left(1 - \delta + \alpha \bar{\theta}_{i,s} (\Theta K_s)^{\alpha-1}\right) \bar{\varphi}_s, \quad (20)$$

where $\bar{\theta}_{i,s} := \frac{\sum_{\mathcal{P}_i} \theta_{j,s}}{|\mathcal{P}_i|}$ denotes the average productivity of firms on investor i 's participation set.

The log return on the investor's portfolio is denoted by $r_{i,s}^p := \log R_{i,s}^p$ and can be written as:

$$r_{i,s}^p = \log(1 + \psi_s(\bar{\theta}_{i,s} - \Theta)) + \log(\bar{R}_s^a \bar{\varphi}_s), \quad (21)$$

where $\psi_s := \frac{\alpha \Theta^\alpha K_s^{\alpha-1}}{1-\delta+\alpha \Theta^\alpha K_s^{\alpha-1}} \frac{1}{\Theta}$ and $\bar{R}_s^a := \frac{1}{N} \sum_{j \in \mathcal{J}} R_{s,j}^a$.

The first term on the right-hand side of Eq. (21) denotes the source of idiosyncratic risk, while the second term represents the log-return on an aggregate (market) portfolio. An important implication of this expression is that the loading of returns on idiosyncratic risk is state-dependent and endogenous, as it depends on the aggregate capital stock in the economy. The amount of idiosyncratic risk born by an investor depends not only on her ability to diversify risk, but also on the aggregate state of the economy. In particular, the idiosyncratic variance of $r_{i,s}^p$, for small values of σ_θ , is given by:

$$\text{Var}_s[r_{i,s}^p] \approx \underbrace{\phi_u \sigma_\theta^2}_{\text{exogenous component}} \times \underbrace{\psi_s^2}_{\text{endogenous component}}, \quad (22)$$

where $\phi_u := \frac{1}{|\mathcal{P}_i|} + \left(1 - \frac{1}{|\mathcal{P}_i|}\right) \rho \in [0, 1]$.

The idiosyncratic variance of $r_{i,s}^p$ depends on an exogenous component, $\phi_u \sigma_\theta^2$, and an endogenous component, ψ_s^2 . The exogenous component depends on firm-level risk σ_θ^2 and the coefficient of under-diversification ϕ_u . When investors have access to all firms, the coefficient of under-diversification is equal to zero, $\phi_u = 0$, and idiosyncratic risk is perfectly diversified. When investors can invest in only a single firm, then $\phi_u = 1$ and there is no diversification. We obtain $0 < \phi_u < 1$ for intermediate levels of market participation, with a higher value of ϕ_u indicating a higher exposure to idiosyncratic risk.

The endogenous component of idiosyncratic variance, ψ_s^2 , is countercyclical, given the procyclical capital stock. 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 counter-cyclical. [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, that shows up not only on stock returns but also on cash-flow (sales) data. These facts can all be explained by our operating-leverage channel, without having to assume shocks to idiosyncratic variance that are correlated across firms.⁹ Moreover, given the endogenous link between idiosyncratic return volatility and aggregate variables, policy interventions can affect the

⁹For the relationship between return risk and operating leverage, see [Lev \(1974\)](#). For evidence on this channel, see [Novy-Marx \(2010\)](#) and [Donangelo et al. \(2019\)](#).

magnitude of idiosyncratic risk in the economy.

2.2.3. 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 using a second-order approximation to the model around an economy without idiosyncratic risk.

Then, we write χ and I as:

$$\chi = \chi^* + \hat{\chi}\sigma_\theta^2 + o(\sigma_\theta^2), \quad I = I^* + \hat{I}\sigma_\theta^2 + o(\sigma_\theta^2). \quad (23)$$

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

In Proposition 1, we describe the impact of idiosyncratic risk on risk-taking and investment. The Appendix provides closed-form expressions for both $\hat{\chi}$ and \hat{I} .

Proposition 1 (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 $\hat{I} > 0$.*

Proof. See Appendix A.3. □

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:

$$0 = \mathbb{E} \left[\mathbb{E}_s \left[C_{i,s}^{-\gamma} R_{j,s}^a \right] \varphi_s^e \right], \quad (24)$$

where $\varphi_s^e \equiv \varphi_s^1 - \varphi_s^0$ is the excess return on the risky technology.

The conditional expectation above can be written as:

$$\mathbb{E}_s \left[C_{i,s}^{-\gamma} R_{j,s}^a \right] \approx \bar{C}_s^{-\gamma} \bar{R}_s^a \times \exp \left(\frac{\gamma(\gamma-1)}{2} \psi_s^2 \phi_u \sigma_\theta^2 \right). \quad (25)$$

The conditional expectation in Eq. (25) acts as a pricing kernel for the investment payoff φ_s^e and has two components. The first component, $\bar{C}_s^{-\gamma} \bar{R}_s^a$, 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 $\gamma > 1$, increases with the amount of idiosyncratic risk investors effectively bear, $\phi_u \psi_s^2 \sigma_\theta^2$.

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 countercyclical consumption risk plays an important role in the determination of risk premia. Yet, 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.

Because the idiosyncratic return risk is countercyclical, the pricing kernel is particularly high in bad times in the case of under-diversification, 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 to 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 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 social planner should intervene in this economy.

2.2.4. Interest rate and the idiosyncratic risk premium

Define the log-SDF for investor i as:

$$m_{i,s} = \log \beta - \gamma(c_{i,s} - c_0), \quad (26)$$

where $c_{i,s} := \log C_{s,i}$ and $c_0 := \log C_0$.

Given the SDF, we can compute the (shadow) riskless rate $r_f := -\log \mathbb{E}[e^{m_{i,s}}]$. Up to second-

order terms, the interest rate is given by the standard expression,

$$r_f = -\log \beta + \gamma \left(\mathbb{E}[c_{i,s}] + \frac{\sigma_c^2}{2} - c_0 \right) - \frac{\gamma(\gamma+1)}{2} (Var[\bar{c}_s] + \phi_u \sigma_\theta^2 \mathbb{E}[\psi_s^2]), \quad (27)$$

where $\bar{c}_s := \mathbb{E}_s[c_{i,s}]$ and $\sigma_c^2 := Var[c_{i,s}^2]$.

The risk-free interest rate depends on the degree of impatience, the intertemporal substitution channel, and a precautionary savings term, which is a function of both aggregate and idiosyncratic risk. Given the level of aggregate risk-taking and investment in the economy, an increase in idiosyncratic risk depresses the interest rate. However, in our production economy, idiosyncratic risk also affects the interest rate through its impact on expect growth and aggregate volatility through the investment decisions.

Let $r_{s,j} \equiv \log R_{j,s}$ denote the log-return on firm j . From the pricing equation for shares (11), we obtain the expected excess return,

$$\mathbb{E}[r_{j,s}] - r_f + \frac{\sigma_r^2}{2} = \gamma Cov(c_{i,s}, r_{j,s}), \quad (28)$$

where σ_r^2 is the variance of the log-returns.

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

$$\mathbb{E}[r_{j,s}] - r_f + \frac{\sigma_r^2}{2} = \underbrace{\gamma Cov(\bar{c}_s, \bar{r}_s)}_{\text{aggregate risk premium}} + \underbrace{\gamma \phi_u \mathbb{E}[\sigma_s^2]}_{\text{idiosyncratic risk premium}}, \quad (29)$$

where $\sigma_s^2 = \psi_s^2 \sigma_\theta^2$ 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 comovement 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 coefficient of under-diversification, ϕ_u . When $\phi_u = 0$, investors are fully diversified and the price of idiosyncratic risk is zero. When $\phi_u = 1$, there is no

diversification 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, the magnitude of idiosyncratic risk is endogenous and potentially affected by economic policy.

3. Idiosyncratic risk externalities

An important aspect of the laissez-faire equilibrium is that the distribution of risk faced by an investor is endogenous, and influenced by both the level and the composition of investment. Firms, however, do not internalize how their investment decisions collectively affect the risk born by others. In this section, we characterize idiosyncratic risk externalities and provide sufficient statistics, based on easily computable risk premia, for the welfare gains achieved by mitigating investment inefficiencies.

3.1. *Assessing constrained efficiency*

We next examine whether the economy is constrained-efficient, that is, whether there are no possible welfare-improving interventions, given the economic constraints.¹⁰ The economy is clearly inefficient as risk is not optimally shared by investors, so a social 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 can improve welfare. For example, could a planner improve welfare by simply altering the investment decisions of firms?

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. We defer the discussion of the implementation of these investment outcomes through financial regulation to Section 5.

¹⁰Constrained efficiency has been the standard way to assess the welfare properties 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).

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. 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, while harming investors. Instead, we are interested in whether there are net gains after the winners of the intervention compensate the 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 investment takes the following form:

$$I^0(\Delta) = I^0 + \kappa_0\Delta, \quad I^1(\Delta) = I^1 + \kappa_1\Delta, \quad (30)$$

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

$$K_s(\Delta) = K_s + (\kappa_0 + \kappa_1\varphi_s^1)N\Delta. \quad (31)$$

Note 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 workers' consumption at the laissez-faire level solves:

$$\bar{C}_{w,s} = (1 - \alpha)(\Theta K_s(\Delta))^\alpha - \tau_s^k(\Delta) K_s(\Delta), \quad (32)$$

where $\bar{C}_{w,s}$ denotes workers' consumption in the laissez-faire equilibrium.

The investor's welfare is then given by:

$$V(\Delta) = u \left(E_0 - \sum_{k=0}^1 I^k(\Delta) \right) + \beta \mathbb{E} [u(C_{i,s}(\Delta))], \quad (33)$$

where, given $R_{i,s}^a(\Delta) = 1 - \delta + \alpha \bar{\theta}_{i,s} \Theta^{\alpha-1} K_s^{\alpha-1}(\Delta)$, consumption at state s can be expressed as

follows:

$$C_{i,s}(\Delta) = \frac{1}{N} [R_{i,s}^a(\Delta)K_s(\Delta) + (1 - \alpha)(\Theta K_s(\Delta))^\alpha - \bar{C}_{w,s}]. \quad (34)$$

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

Our first main result, Proposition 2, demonstrates that investment is inefficiently low in the laissez-faire equilibrium. For that, we consider a perturbation that increases the expected value of capital per capita by Δ , while keeping the variance of K_s constant, that is, we set $\kappa_0 = 1$ and $\kappa_1 = 0$.

Proposition 2. *Suppose $\kappa_0 = 1$ and $\kappa_1 = 0$. The marginal gain of increasing Δ , in terms of initial consumption, is given by:*

$$\frac{V'(0)}{u'(C_0)} = -(1 - \alpha)\mathbb{E} [\text{Cov}_s (M_{i,s}, R_{i,s}^a)] > 0. \quad (35)$$

This marginal gain, up to second-order in the amount of idiosyncratic risk, is given by:

$$\frac{V'(0)}{u'(C_0)} \approx (1 - \alpha) \left[\underbrace{\gamma\phi_u \mathbb{E} [\sigma_s^2]}_{\text{id. risk premium}} + \gamma\phi_u \underbrace{\left(\mathbb{E}^{\mathbb{Q}} [\sigma_s^2] - \mathbb{E} [\sigma_s^2] \right)}_{\text{id. variance risk premium}} \right]. \quad (36)$$

Moreover, the idiosyncratic variance risk premium is positive, i.e., $\mathbb{E}^{\mathbb{Q}} [\sigma_s^2] - \mathbb{E} [\sigma_s^2] > 0$.¹¹

Proof. See Appendix B.1. □

Proposition 2 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.

¹¹The risk-neutral probabilities satisfy $\mathbb{E}^{\mathbb{Q}}[X_s] = \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a X_s \right]$ for all random variables X_s , using $\mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \right] = 1$.

An increase in the 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 Eq. (22). 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 call them idiosyncratic risk externalities.

As seen in Subection 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 increase investment in the second-best compared with what occurs in the laissez-faire economy. This is because firms do not internalize a potential benefit from 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, yet, a further increase in investment achieves a welfare gain. This is a typical feature of second-best applications.

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 Q-measure is the relevant one with which to compute expected welfare gains. By definition, one dollar in a high-probability state under the 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 was constant across states, this distinction would not be necessary, but given the countercyclicity of return risk, important deviations between the physical and the risk-neutral measures of expected variance can occur. In particular, because the idiosyncratic variance is larger in high marginal utility states, ex-

pected variance is higher under the risk-neutral measure, implying a positive idiosyncratic variance risk premium. Note 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 degree of diversification.

An alternative way to write Eq. (36) is $(1 - \alpha)\gamma\phi_u\mathbb{E}^Q[\sigma_s^2]$, 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. 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 Eq. (36) is that it depends on the labor share $1 - \alpha$. For instance, 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, as formally stated below.

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.*

The return risk is completely exogenous when $\alpha = 1$, as can be seen from Eq. (22), given that ψ_s is constant when $\alpha = 1$. 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: endogenous return risk and under-diversification. It is the interaction of these two ingredients that opens the door to welfare-improving interventions.¹²

3.3. Excessive aggregate risk-taking

Our second main result, Proposition 3, demonstrates that investment is excessively risky in the laissez-faire equilibrium. For that, we consider an intervention that reduces the share invested

¹²The 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).

in the risky technology. In particular, we choose κ_0 and κ_1 such that the risk-neutral standard deviation of capital per capita decreases by Δ , while we keep the total investment unchanged.

Proposition 3. *Suppose $\kappa_0 = \frac{1}{\sqrt{\text{Var}^{\mathbb{Q}}[\varphi^1]}}$ and $\kappa_1 = -\frac{1}{\sqrt{\text{Var}^{\mathbb{Q}}[\varphi^1]}}$. The marginal gain of increasing Δ , in terms of date $t = 0$ consumption, is given by:*

$$\frac{V'(0)}{u'(C_0)} = (1 - \alpha) \mathbb{E} [\text{Cov}_s (M_{i,s}, R_{i,s}) \varphi_s^e] \kappa_0 > 0. \quad (37)$$

This marginal gain, up to second-order in the amount of idiosyncratic risk, is given by:

$$\begin{aligned} \frac{V'(0)}{u'(C_0)} &\approx (1 - \alpha) \gamma \phi_u \text{Cov}^{\mathbb{Q}}(\sigma_s^2, \varphi_s^e) \kappa_1 \\ &= (1 - \alpha) \gamma \phi_u \underbrace{\left(\mathbb{E}^{\mathbb{Q}} [\sigma_s^2] - \mathbb{E} [\sigma_s^2] \right)}_{\text{id. variance risk premium}} \zeta > 0, \end{aligned} \quad (38)$$

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

Proof. See Appendix B.1. □

Proposition 3 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, that is, $\text{Cov}^{\mathbb{Q}}(\sigma_s^2, \varphi_s^e) < 0$. 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 discussed in Section 2. 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 directly exposed 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 countercyclical 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

We consider several extensions that evaluate the robustness of our main results and offer generalizations in Supplement OS.1, available online.

Intermediate inputs and endogenous labor supply. 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 source of variation in 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 version derived for 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 show that an economy with an elastic labor supply is equivalent to an economy with an elastically supplied intermediate good. Therefore, a higher labor supply elasticity dampens the cyclicity of margins and idiosyncratic risk externalities.

The role of financial market participation. In our baseline model, we show that the economy is constrained inefficient under the assumption that workers do not participate in financial markets. We next show that our results do not hinge on this particular assumption.

In the context of the model with intermediate goods, we allow for an arbitrary distribution of profits from intermediate-good firms between investors and hand-to-mouth entrepreneurs. In particular, we allow for the limit case in which investors receive all profits, and there are no hand-to-mouth agents. We show that the characterization of idiosyncratic risk externalities does not depend on the underlying ownership shares. Given the equivalence between an economy with intermediate-good production and an economy with labor, idiosyncratic risk externalities are still present if investors receive all the labor income. This highlights that these externalities do not hinge on the limited financial market participation of workers, but on endogenous return disper-

sion.

Substitution patterns between labor and capital. Next, we 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 tend to amplify the effects we characterize. In the empirical literature, values for this elasticity are typically below one. [Oberfield and Raval \(2021\)](#) report an elasticity in the range of 0.5-0.7, which suggests that the gains from the proposed intervention may be actually higher than what the baseline Cobb-Douglas case indicates.

Epstein-Zin preferences. We extend the model to allow for Epstein-Zin preferences, which allow us to disentangle the role of risk aversion and the elasticity of intertemporal substitution (EIS). We find that Eqs. (35) and (37) hold in this case, where $M_{i,s}$ is now the SDF induced by Epstein-Zin preferences. The small risk approximations in Eqs. (36) and (38) also hold in this case, showing that it is the risk aversion coefficient γ that is relevant for the determination of the externality instead of the (inverse) of the EIS. We also show how it is the EIS that controls whether investment in the economy with idiosyncratic risk is larger or smaller than at the laissez-faire level.

Endogenous diversification. 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 cost 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. The welfare impact of idiosyncratic risk externalities

We have shown that idiosyncratic risk can cause inefficiencies when agents hold under-diversified portfolios. But it remains unclear whether these inefficiencies are quantitatively important enough to warrant policy interventions. In this section, we quantify the significance of the idiosyncratic risk externalities. We evaluate the welfare cost of underinvestment by measuring expression (36), which involves the estimates of the idiosyncratic risk premium, the idiosyncratic variance risk premium, and the labor share. We also assess the welfare cost of excessive risk-taking by measuring expression (38), which requires an estimate of the risk-neutral probabilities. Note that our quantification procedure exploits asset-market data, embodying the idea that policy decisions can be informed by inputs from financial markets.

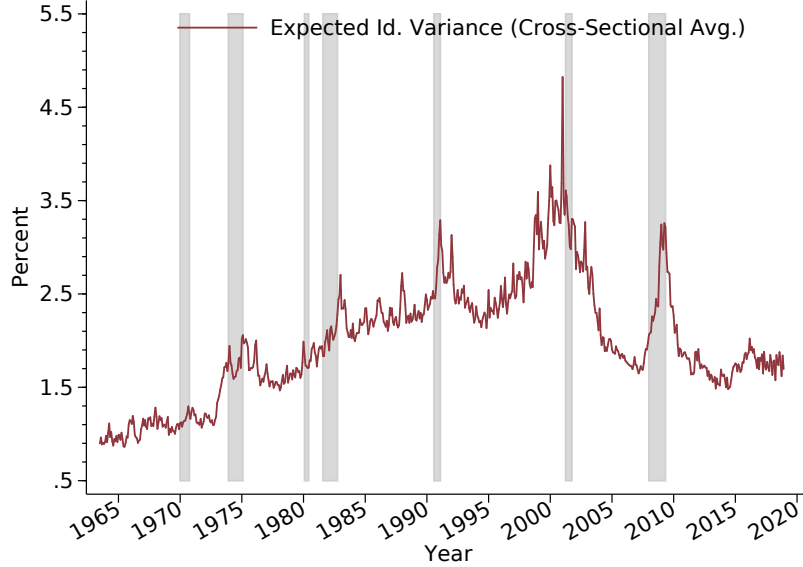
4.1. Parameter choices

We next present the estimates of the aforementioned quantities, which underlie our parameter choices.

Estimate of the idiosyncratic risk premium. The idiosyncratic risk premium is the product of the magnitude and the price of idiosyncratic risk. We compute a measure of the magnitude of risk using an EGARCH model, following Fu (2009). Specifically, within a comprehensive sample of U.S. stocks, we estimate, stock by stock, an augmented Fama-French three-factor model that accounts for heteroskedasticity. For each stock, the estimation produces a monthly series of the conditional variance of residuals, which we adopt as our measure of idiosyncratic risk for this stock. In Fig. 2, we plot the cross-sectional average of this measure (which we will refer to as “idiosyncratic variance” hereafter). One can clearly see countercyclicality in this series, as there are sizeable spikes in almost every recession. The mean and median of the idiosyncratic variance in our sample is 1.84% and 0.94%, respectively (that is, 22.1% and 11.3%, respectively, in annualized terms). Based on these numbers, we set $\mathbb{E}[\sigma_s^2]$ equal to 11.3%.

For an estimate of the price of idiosyncratic risk, we employ Fama-MacBeth regressions. That is, we regress excess stock returns on our measure of idiosyncratic risk, along with several other salient characteristics like size, book-to-market ratio, and so on. We provide the estimation details

Fig. 2. Idiosyncratic variance averaged across stocks. This figure displays, month by month, the cross-sectional average of the (expected) idiosyncratic variance, which is estimated by fitting an EGARCH model to individual U.S. stocks. Shaded areas indicate recessions in the U.S. as defined by National Bureau of Economic Research (NBER).



in Appendix C, where we also report some robustness tests. We find that idiosyncratic risk has strong explanatory power for average returns: a one percentage point increase in the idiosyncratic variance is associated with a 35-52 bps increase in average return. This leads us to set $\gamma\phi_u = 0.35$.

Estimate of the idiosyncratic variance risk premium. Besides the idiosyncratic risk premium, we also need an estimate of the idiosyncratic variance risk premium. Even though the idiosyncratic variance risk premium has not been directly considered in the literature, we show that it is possible to recover its value from estimates of the firm-level and market-level variance risk premia. First, based on the return decomposition proposed by [Campbell et al. \(2001\)](#), we can write the average return variance as follows:

$$\bar{\sigma}_t^2 = \sigma_{m,t}^2 + \bar{\sigma}_{id,t}^2 \quad (39)$$

where $\bar{\sigma}_t^2$ is the average of the return variance across individual stocks; $\sigma_{m,t}^2$ is the market variance; and $\bar{\sigma}_{id,t}^2$ is the cross-sectional average of the idiosyncratic variance. Details of this calculation are provided in Appendix C.2. Then from the definition of the idiosyncratic variance risk premium

(that is, $\overline{VRP}_{id,t} \equiv \mathbb{E}_t^Q[\overline{\sigma}_{id,t+1}^2] - \mathbb{E}_t[\overline{\sigma}_{id,t+1}^2]$), we immediately get:

$$\overline{VRP}_{id,t} = \underbrace{\mathbb{E}_t^Q[\overline{\sigma}_{t+1}^2] - \mathbb{E}_t[\overline{\sigma}_{t+1}^2]}_{\overline{VRP}_t} - \underbrace{(\mathbb{E}_t^Q[\sigma_{m,t+1}^2] - \mathbb{E}_t[\sigma_{m,t+1}^2])}_{VRP_{m,t}}, \quad (40)$$

which implies that the (average) idiosyncratic variance risk premium can be computed by subtracting the market variance risk premium from the average variance risk premium for individual stocks. The latter two quantities have been estimated by [Han and Zhou \(2012\)](#), who report that \overline{VRP}_t in their sample is 5.88% (in annualized terms), while the $VRP_{m,t}$ is 2.83%. Therefore, we let $\mathbb{E}^Q[\sigma_s^2] - \mathbb{E}[\sigma_s^2]$ in our model be equal to $5.88\% - 2.83\% \approx 3.05\%$.

Estimate of the labor share. We set $\alpha = 0.33$ to match the share of labor income to total income, which is about 66% in the U.S. aggregate data, a standard choice for this parameter.

Estimate of the risk-neutral probabilities. Lastly, to measure Eq. (38), we also need to quantify the physical and the risk-neutral probabilities of the aggregate states. To this end, we define high (low) states as periods when the idiosyncratic variance averaged across stocks is below (above) the median. Thus by definition, the physical probability of high (low) states is 0.5 (that is, $p_l = 0.5$). As for the risk-neutral probabilities, we impute the value from the variance risk premium. Specifically, we use the following expression for the variance risk premium:

$$VRP = (q_l - p_l)(\sigma_l^2 - \sigma_h^2) \quad (41)$$

to compute q_l from VRP , given p_l as well as σ_s^2 , the idiosyncratic variance in state $s \in \{low, high\}$. Given the estimate of $VRP = 2.83\%$ mentioned above, we set q_l to 0.75.

4.2. Welfare analysis

As we have pinned down all the parameters needed to evaluate Eqs. (36) and (38), we next move on to quantify the welfare impact of the idiosyncratic risk externalities. Our approach features a marginal analysis in which we assess the welfare gains of small changes in the laissez-faire investment policy. We find substantial welfare gains from increasing investment levels and reducing risk-taking using our sufficient-statistic approach.

Under-investment. First, we consider the welfare effects of an increase in investment in the riskless technology. We have demonstrated that the corresponding change in welfare is given by expression (36), which can be rewritten as $IRE_I = (1 - \alpha) [IRP + \gamma\phi_u VRP]$, where IRP denotes the idiosyncratic risk premium and VRP the idiosyncratic variance risk premium. Given our parameter choices presented above, we compute the value of IRE_I to be $\frac{2}{3} \times (0.35 \times 11.3\% + 0.35 \times 3.0\%) \approx 3.3\%$. This means that, for an extra dollar of investment on top of the amount chosen by agents, there is a welfare gain of 3.3 cents (in current-consumption terms) that is not privately taken into account. This is a nontrivial magnitude. For concreteness, consider the value of the capital stock in the U.S. economy, which is around \$50 trillion. An IRE_I of 3.3% means that an extra 1% increase in the capital stock (that is, about \$560 billion) produces a welfare gain of \$16.8 billion that is not privately taken into account.

Excessive risk-taking. We also consider the welfare effects of a reallocation of investment, moving resources from the risky technology to the riskless technology. That would cause a welfare change of magnitude represented by Eq. (38). That equation can be rewritten as $IRE_\chi = (1 - \alpha)\gamma\phi_u \cdot VRP \cdot \zeta$, where $\zeta = \frac{\sqrt{q_h q_l}}{q_l - p_l}$ is determined by the physical and the risk-neutral probabilities of the aggregate states. Substituting in our parameter values, we obtain $IRE_\chi = \frac{2}{3} \times 0.35 \times 3.0\% \times 1.7 \approx 1.2\%$. This means that, at the margin, shifting one dollar of investment away from the risky technology to the riskless technology can bring an uninternalized welfare gain equivalent to 1.2 cents in current-consumption terms. This is again a sizable improvement.

In sum, we show that welfare improvements from deviating from privately optimal investment policies towards higher investment levels and lower risk taking can be meaningful, given market data on risk premia and standard parameter values. In Supplement OS.2, we provide an extended discussion and additional interpretations of those welfare gains. In Section 5, we discuss how financial regulation can be designed to take advantage of those potential improvements.

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 quantitative magnitude of these frictions from an unconditional perspective. As the importance

of the frictions may vary with the state of the economy, we next consider a dynamic extension of our sufficient statistic formulas. Our goal is not to provide the most general dynamic model, but instead to offer a minimal deviation from the environment we have considered so far. For this reason, we consider an overlapping generations version of the baseline two-period model described in Section 2, which will allow us to quantify how the idiosyncratic risk externalities evolve over time.

Dynamic model. The economy is populated by a finite number 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 as a snapshot of the dynamic economy just described. 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 interventions studied for the baseline model, in Section 3, we can consider a change in investment in period t that keeps the income of the next generation constant. Hence, the new generation plays the role that workers played in the baseline model. In Proposition 4, we characterize the risk externalities in this dynamic model.

Proposition 4 (Conditional risk externalities). *Consider the effects of regulating investment decisions in the dynamic economy. Then,*

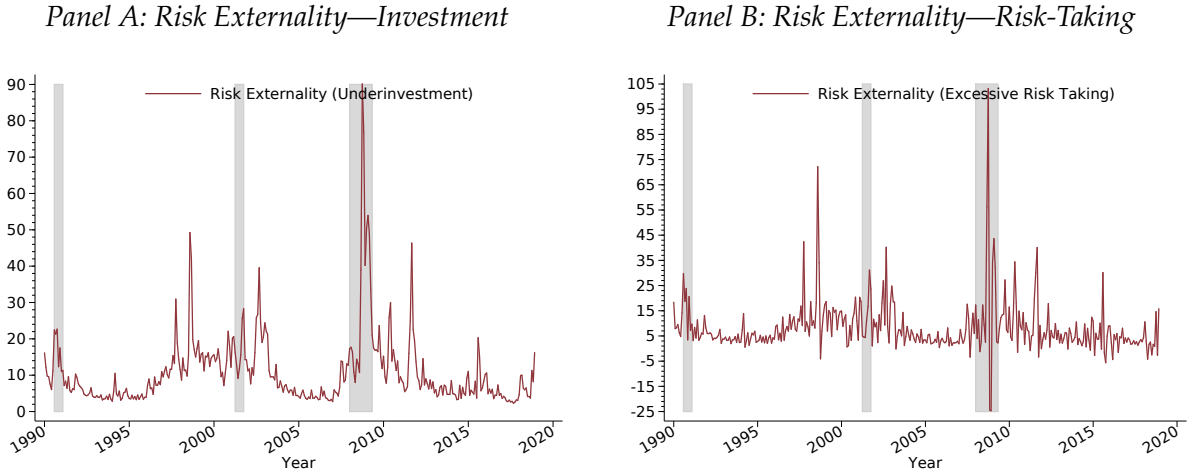
i. Investment

$$\frac{V'_s(0)}{u'(C_{s,0})} \approx (1 - \alpha) \left[\underbrace{\gamma\phi_u \mathbb{E}_s[\sigma_{s'}^2]}_{id. \text{ risk premium}} + \gamma\phi_u \underbrace{\left(\mathbb{E}_s^Q[\sigma_{s'}^2] - \mathbb{E}_s[\sigma_{s'}^2] \right)}_{id. \text{ variance risk premium}} \right]. \quad (42)$$

ii. Aggregate risk-taking

$$\frac{V'_s(0)}{u'(C_{s,0})} \approx (1 - \alpha) \gamma\phi_u \underbrace{\left(\mathbb{E}_s^Q[\sigma_{s'}^2] - \mathbb{E}_s[\sigma_{s'}^2] \right)}_{id. \text{ variance risk premium}} \zeta_s > 0, \quad (43)$$

Fig. 3. 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.



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 $s \in \mathcal{S}$.

The expressions comprising Proposition 4 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 economic efficiency. 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 Fig. 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.

Fig. 3 shows 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 more severe in bad times, when idiosyncratic uncertainty is high. In particular, the two externality measures spiked during recent financial crises, indicating that those were periods in which the disagreement between the social planner and the private agents' investment and risk-taking incentives was highest.

The time-variation 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 the countercyclical capital buffer (CCyB) included in Basel III.

5. Risk externalities and financial regulation

We next address two questions related to the regulation of risk externalities: implementation and optimal regulation. When deriving the sufficient statistics for the externalities discussed in Section 3, we assume that the planner can 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 benefit on debt and risk-weighted capital requirements, are capable of implementing the desired allocation. We also show how asset-price data can be used to determine the optimal levels of these instruments.

In Appendix D.1, we introduce financial intermediaries and regulation in an extension of the baseline model. Intermediaries raise funds from investors, issuing a mix of debt and equity, and allocate these funds to non-financial firms. Importantly, these intermediaries are subject to regulatory constraints.

A tax benefit on debt, τ_d , generates a cost advantage in its issuance. In isolation, this advantage would make intermediaries favor debt as their only source of financing. Additionally, in the design of regulation, the tax benefit can be used to determine the intermediaries' cost of funds and, through that, it can influence the investment level of the economy.

However, financial intermediaries are also subject to a risk-weighted capital requirement constraint of the following form:

$$\sum_k I_j^k - P_d D_j \geq \sum_k \omega^k I_j^k, \quad (44)$$

where $D_j > 0$ is the level of debt raised and P_d is its equilibrium price. When they raise debt instead of equity, it tightens this constraint, counterbalancing its tax advantage. Therefore, intermediaries face a meaningful capital structure trade-off. Moreover, the particular risk-weights, ω^k , imposed on assets of class $k \in \{0, 1\}$, can serve to relatively discourage investment in that class and, through that, control investment composition in the economy.

We show in Supplement OS.5 that any allocation with the following properties can be implemented through appropriately chosen tax benefit of debt and risk weights: the allocation (i) is feasible, (ii) is constrained by limited participation, (iii) features implicit subsidies to investment and (iv) features implicit taxes on risk-taking.

Investment and risk-taking wedges. Consider the capital structure choice of the financial intermediary. We show in Appendix D that, given the tax benefit ($\tau_d \geq 0$), the level of debt ($D_j > 0$) and its equilibrium price (P_d), the following distortion in the Euler equation for the overall investment level, I :

$$\mathbb{E} \left[M_{j,s} R_{j,s}^a \frac{K_{s,j}}{I - P_d \tau_d D_j} \right] = 1. \quad (45)$$

The tax benefit essentially reduces the cost of investment, creating a wedge in the investment Euler equation. Notice that risk weights do not directly affect this equation. In contrast, they have a direct impact on the Euler equation for the investment composition:

$$\mathbb{E} \left[M_{j,s} R_{j,s}^a \varphi_s^e \right] = (\omega^1 - \omega^0) P_d \tau_d. \quad (46)$$

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 change the covariance of φ_s^e with the SDF, such that it matches the right-hand side of Eq. (46). The term $P_d \tau_d$ captures the shadow cost of the regulatory constraint. The intermediary optimally balances this shadow cost with the tax benefit.

We then solve for the optimal policy and show that it requires wedges of the following form:

$$P_d \tau_d \frac{D_j}{I} = IRE_I \quad (47)$$

and

$$(\omega^1 - \omega^0) \tau_d P_d = IRE_\chi, \quad (48)$$

where $IRE_I \approx (1 - \alpha) \gamma \phi_u [(1 - \chi) \mathbb{E}^Q[\sigma_s^2] + \chi \mathbb{E}^Q[\sigma_s^2 \varphi_s^1]]$ and $IRE_\chi \approx -(1 - \alpha) \gamma \phi_u Cov^Q(\sigma_s^2, \varphi_s^e)$ are, respectively, measures of the marginal impact of the investment level and of the investment composition on the idiosyncratic risk externalities. These measures are analogous to those presented in Propositions 2 and 3. Intuitively, unlike individual investors under *laissez-faire*, the social planner internalizes the impact of investment decisions on the distribution of idiosyncratic risk.

Eq. (47) shows that the tax benefit, per unit of investment, should equal the risk externality on investment. All the elements required to estimate the tax benefit can be recovered directly from

the data, as illustrated in Section 4. This sufficient-statistic approach contrasts with alternative frameworks for financial regulation design, which typically rely on the calibration and numerical solution of an economic model.

Eq. (48) connects the risk weights, the tax benefit on debt, and the risk externality on aggregate risk-taking. Its 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 invested in the risky technology tightens the regulatory constraint, while $\tau_d P_d$ 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.

A valuable property of Eq. (48) for the determination of risk weights is again that it can be estimated directly from the data. In our empirical exercise in Section 4, we connect 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, providing a tight connection between the data on asset prices and the optimal regulatory risk weights.¹³

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 underdiversified 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, which can be measured directly in the data. In particular, these statistics are derived from two risk premia: an idiosyncratic risk premium and an

¹³In particular, the risk weight on asset k is given by $\omega^k - \omega^0 = \kappa \text{Cov}^Q(\sigma_s^2, \varphi_s^{e,k})$, given a positive constant $\kappa > 0$, for $k \in \{1, 2, \dots, K\}$ and K is the number of risky assets.

idiosyncratic variance risk premium.

We find that idiosyncratic uncertainty has a significant impact on welfare and also consider the optimal financial regulation. The optimal allocation can be implemented using two instruments: a tax benefit on debt and risk-weighted capital requirements on financial intermediaries. Intuitively, the tax benefit 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 further extended in several other directions. 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 find. Additionally, the financial intermediaries we consider are not subject to any frictions other than regulation itself. 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 intermediary asset-pricing literature (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 relevant to the design of financial regulation.

References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., Tahbaz-Salehi, A., 2012. The network origins of aggregate fluctuations. *Econometrica* 80, 1977–2016.
- Admati, A. R., Pfleiderer, P., Zechner, J., 1994. Large shareholder activism, risk sharing, and financial market equilibrium. *Journal of Political Economy* 102, 1097–1130.
- Ang, A., Hodrick, R. J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *The Journal of Finance* 61, 259–299.
- Angeletos, G.-M., Calvet, L.-E., 2006. Idiosyncratic production risk, growth and the business cycle. *Journal of Monetary Economics* 53, 1095–1115.
- Bekaert, G., Hodrick, R. J., Zhang, X., 2012. Aggregate idiosyncratic volatility. *Journal of Financial and Quantitative Analysis* 47, 1155–1185.
- Bloom, N., 2009. The impact of uncertainty shocks. *Econometrica* 77, 623–685.
- Blume, M. E., Friend, I., 1975. The asset structure of individual portfolios and some implications for utility functions. *The Journal of Finance* 30, 585–603.
- Bocola, L., Lorenzoni, G., 2020. Risk sharing externalities. Working paper, National Bureau of Economic Research.
- Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. *The Review of Financial Studies* 22, 4463–4492.
- Brav, A., Constantinides, G. M., Geczy, C. C., 2002. Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy* 110, 793–824.
- Brody, D. C., Hughston, L. P., Macrina, A., 2008. Dam rain and cumulative gain. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 464, 1801–1822.
- Brunnermeier, M. K., Sannikov, Y., 2014. A macroeconomic model with a financial sector. *American Economic Review* 104, 379–421.

- Caballero, R. J., Lorenzoni, G., 2014. Persistent appreciations and overshooting: A normative analysis. *IMF Economic Review* 62, 1–47.
- Calvet, L. E., Campbell, J. Y., Sodini, P., 2007. Down or out: assessing the welfare costs of household investment mistakes. *Journal of Political Economy* 115, 707–747.
- Campbell, J. Y., Lettau, M., Malkiel, B. G., Xu, Y., 2001. Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *The Journal of Finance* 56, 1–43.
- Chen, H., Miao, J., Wang, N., 2010. Entrepreneurial finance and nondiversifiable risk. *The Review of Financial Studies* 23, 4348–4388.
- Chen, Z., Strebulaev, I. A., 2018. Macroeconomic risk and idiosyncratic risk-taking. *The Review of Financial Studies* 32, 1148–1187.
- Chordia, T., Subrahmanyam, A., Anshuman, V., 2001. Trading activity and expected stock returns. *Journal of Financial Economics* 59, 3 – 32.
- Christiano, L. J., Motto, R., Rostagno, M., 2014. Risk shocks. *American Economic Review* 104, 27–65.
- Constantinides, G. M., Duffie, D., 1996. Asset pricing with heterogeneous consumers. *Journal of Political Economy* 104, 219–240.
- Constantinides, G. M., Ghosh, A., 2017. Asset pricing with countercyclical household consumption risk. *The Journal of Finance* 72, 415–460.
- Dávila, E., Korinek, A., 2018. Pecuniary externalities in economies with financial frictions. *The Review of Economic Studies* 85, 352–395.
- DeMarzo, P. M., Urošević, B., 2006. Ownership dynamics and asset pricing with a large shareholder. *Journal of Political Economy* 114, 774–815.
- Di Tella, S., 2017. Uncertainty shocks and balance sheet recessions. *Journal of Political Economy* 125, 2038–2081.
- Di Tella, S., 2019. Optimal regulation of financial intermediaries. *American Economic Review* 109, 271–313.

- Donangelo, A., Gourio, F., Kehrig, M., Palacios, M., 2019. The cross-section of labor leverage and equity returns. *Journal of Financial Economics* 132, 497–518.
- Dou, W. W., 2016. Embrace or fear uncertainty: growth options, limited risk sharing, and asset prices. Working Paper. Massachusetts Institute of Technology .
- Drechsler, I., Yaron, A., 2010. What's vol got to do with it. *The Review of Financial Studies* 24, 1–45.
- Eiling, E., 2013. Industry-specific human capital, idiosyncratic risk, and the cross-section of expected stock returns. *The Journal of Finance* 68, 43–84.
- Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. *The Journal of Finance* 47, 427–465.
- Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3 – 56.
- Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607–636.
- Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91, 24 – 37.
- Gabaix, X., 2011. The granular origins of aggregate fluctuations. *Econometrica* 79, 733–772.
- Gârleanu, N., Panageas, S., 2018. Finance in a time of disruptive growth, Working Paper. University of California, Los Angeles.
- Gârleanu, N., Panageas, S., Yu, J., 2015. Financial entanglement: a theory of incomplete integration, leverage, crashes, and contagion. *American Economic Review* 105, 1979–2010.
- Gârleanu, N., Panageas, S., Yu, J., 2020. Impediments to financial trade: Theory and applications. *The Review of Financial Studies* 33, 2697–2727.
- Geanakoplos, J. D., Polemarchakis, H. M., 1986. Existence, regularity, and constrained suboptimality of competitive allocations. *Essays in Honor of Kenneth J. Arrow: Volume 3, Uncertainty, Information, and Communication* 3, 65.

- Gomes, J. F., 2001. Financing investment. *American Economic Review* 91, 1263–1285.
- Greenwald, B. C., Stiglitz, J. E., 1986. Externalities in economies with imperfect information and incomplete markets. *The Quarterly Journal of Economics* 101, 229–264.
- Gromb, D., Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics* 66, 361–407.
- Han, B., Zhou, Y., 2012. Variance risk premium and cross-section of stock returns. Unpublished paper, University of Texas at Austin .
- Hart, O. D., 1975. On the optimality of equilibrium when the market structure is incomplete. *Journal of Economic Theory* 11, 418–443.
- He, Z., Kondor, P., 2016. Inefficient investment waves. *Econometrica* 84, 735–780.
- He, Z., Krishnamurthy, A., 2013. Intermediary asset pricing. *American Economic Review* 103, 732–70.
- Herskovic, B., Kelly, B., Lustig, H., Van Nieuwerburgh, S., 2016. The common factor in idiosyncratic volatility: quantitative asset pricing implications. *Journal of Financial Economics* 119, 249–283.
- Himmelberg, C. P., Hubbard, R. G., Love, I., 2000. Investment, protection, ownership, and the cost of capital. National Bank of Belgium Working Paper .
- Hirshleifer, D., 1988. Residual risk, trading costs, and commodity futures risk premia. *The Review of Financial Studies* 1, 173–193.
- Hoyle, E., 2010. Information-based models for finance and insurance. arXiv preprint arXiv:1010.0829 .
- Iachan, F. S., 2020. Capital budgeting and risk taking under credit constraints. *Management Science* 66, 4292–4314.
- Itskhoki, O., Moll, B., 2019. Optimal development policies with financial frictions. *Econometrica* 87, 139–173.

- Ivković, Z., Weisbenner, S., 2005. Local does as local is: information content of the geography of individual investors' common stock investments. *The Journal of Finance* 60, 267–306.
- Jeanne, O., Korinek, A., 2019. Managing credit booms and busts: A pigouvian taxation approach. *Journal of Monetary Economics* 107, 2–17.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *The Journal of Finance* 48, 65–91.
- Kelly, M., 1995. All their eggs in one basket: Portfolio diversification of us households. *Journal of Economic Behavior & Organization* 27, 87–96.
- Khorrami, P., 2019. The risk of risk-sharing: Diversification and boom-bust cycles, Working Paper. Imperial College Business School.
- Lanteri, A., Rampini, A. A., 2020. Constrained-efficient capital reallocation, Working Paper. Duke University.
- Lev, B., 1974. On the association between operating leverage and risk. *Journal of Financial and Quantitative Analysis* 9, 627–641.
- Levy, H., 1978. Equilibrium in an imperfect market: a constraint on the number of securities in the portfolio. *The American Economic Review* 68, 643–658.
- Lorenzoni, G., 2008. Inefficient credit booms. *The Review of Economic Studies* 75, 809–833.
- Mehra, R., Wahal, S., Xie, D., 2019. Is idiosyncratic risk conditionally priced? Tech. rep., National Bureau of Economic Research, Inc.
- Merton, R. C., 1987. A simple model of capital market equilibrium with incomplete information. *The Journal of Finance* 42, 483–510.
- Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Novy-Marx, R., 2010. Operating leverage. *Review of Finance* 15, 103–134.
- Oberfield, E., Raval, D., 2021. Micro data and macro technology. *Econometrica* 89, 703–732.

- Panousi, V., Papanikolaou, D., 2012. Investment, idiosyncratic risk, and ownership. *The Journal of Finance* 67, 1113–1148.
- Polkovnichenko, V., 2005. Household portfolio diversification: a case for rank-dependent preferences. *The Review of Financial Studies* 18, 1467–1502.
- Schmitt-Grohé, S., Uribe, M., 2004. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control* 28, 755–775.
- Silva, D., Townsend, R., 2019. Risk-taking over the life cycle: aggregate and distributive implications of entrepreneurial risk, Working Paper. Massachusetts Institute of Technology.
- Spiegel, M. I., Wang, X., 2007. Cross-sectional variation in stock returns: liquidity and idiosyncratic risk, Working Paper. Yale School of Management.
- Stiglitz, J. E., 1982. The inefficiency of the stock market equilibrium. *The Review of Economic Studies* 49, 241–261.
- Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. *The Review of Economic Studies* 77, 779–805.
- Wang, C., Wang, N., Yang, J., 2012. A unified model of entrepreneurship dynamics. *Journal of Financial Economics* 106, 1–23.
- Welch, I., 2019. Simpler better market betas, Working Paper. University of California, Los Angeles.
- Zhou, H., 2018. Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty. *Annual Review of Financial Economics* 10, 481–497.

Appendix A. Appendix for Section 2

A.1. The gamma process and the limit economy with a continuum of firms

A gamma process is a stochastic process with independent gamma-distributed increments. Let the (shape) parameter ν control the rate of jump arrivals and the (rate) parameter λ inversely control the jump size. We denote the process by $\Gamma(t, \nu, \lambda)$ or Γ_t , with some abuse of notation.

The gamma bridge is obtained by fixing a gamma process on a given position at a particular location t . Take the bridge $\{\Gamma_t^B\}$ that has final value 1 at $t = 1$. It can be constructed from an arbitrary gamma process $\Gamma(t, \nu, \lambda)$ by setting $\Gamma_t^B = \frac{\Gamma_t}{\Gamma_1}$. By the scaling property of the underlying gamma process, its law is invariant to the rate parameter λ (see Hoyle, 2010, p. 26 and Brody et al., 2008).

For an arbitrary number of firms, N , construct $\theta_{j,s} = \Theta N \left[\Gamma_j^B - \Gamma_{j-\frac{1}{N}}^B \right]$. By the properties of the gamma distribution, each disturbance component $\Gamma_j^B - \Gamma_{j-\frac{1}{N}}^B \sim Be\left(\nu \frac{1}{N}, \nu \frac{N-1}{N}\right)$, where $Be(\alpha, \beta)$ denotes the beta distribution with its two shape parameters.¹⁴ Also, the joint distribution of these components is a symmetric Dirichlet distribution and, by construction, TFP risk is idiosyncratic, because $\frac{\sum \theta_j}{N} \stackrel{a.s.}{=} \Theta = \mathbb{E}[\theta_{j,s}]$. Note that, from the properties of the Dirichlet distribution, we have that the correlation between the productivity of any two distinct firms is given by $\rho = -\frac{1}{N-1}$.

Let $\bar{\theta}_{i,s} = \sum_{j \in \mathcal{P}_i} \theta_j / |\mathcal{P}_i|$. Then, $\bar{\theta}_{i,s} = \Theta \frac{x_{i,s}}{\phi}$, where $\phi = \frac{|\mathcal{P}_i|}{N} \in \left\{ \frac{1}{N}, \frac{2}{N}, \dots, 1 \right\}$ represents the effective coverage of the agent's participation set and $x_{i,s} \sim Be(\nu\phi, \nu(1-\phi))$. It follows that the distribution of investor's returns and consumption does not depend on the particular N , only on the coverage of her participation set. This allows us to study a limit economy with a continuum of firms and well-behaved idiosyncratic risk at the investor level.

A.2. Asymptotic analysis

Consider a family of economies parameterized by σ_θ , where the productivity of firm j is given by

$$\theta_{j,s}(\sigma_\theta) = \Theta + \sigma_\theta \epsilon_{j,s}, \quad (\text{A.1})$$

¹⁴Notice that $\Gamma_j^B - \Gamma_{j-\frac{1}{N}}^B = \left(\Gamma_j - \Gamma_{j-\frac{1}{N}} \right) \left\{ \Gamma_j - \Gamma_{j-\frac{1}{N}} + \Gamma_1 - \left(\Gamma_j - \Gamma_{j-\frac{1}{N}} \right) \right\}^{-1}$ represents the proportion between $\Gamma_j - \Gamma_{j-\frac{1}{N}} \sim \Gamma\left(\frac{\nu}{N}, \lambda\right)$ and its complement $\Gamma_1 - \left(\Gamma_j - \Gamma_{j-\frac{1}{N}} \right) \sim \Gamma\left(\nu \frac{N-1}{N}, \lambda\right)$.

where $\{\epsilon_{j,s}\}_{j \in \mathcal{J}}$ is a collection of random variables satisfying $\mathbb{E}[\epsilon_{j,s}] = 0$, $\text{Var}[\epsilon_{j,s}] = 1$, and pairwise correlation $\rho = -\frac{1}{N-1}$.

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

$$C_0(\sigma_\theta) = C_0^* + \hat{C}_0 \sigma_\theta^2 + o(\sigma_\theta^2) \quad (\text{A.2})$$

$$I^k(\sigma_\theta) = I^{k,*} + \hat{I}^k \sigma_\theta^2 + o(\sigma_\theta^2), \quad (\text{A.3})$$

for $k = 0, 1$.

Note that this already imposes the result that the first-order term for C_0 and I_k are equal to zero, in a way analogous to the findings in [Schmitt-Grohé and Uribe \(2004\)](#). 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 C_0 and I^k respond to the presence of idiosyncratic risk, i.e., to solve for the impact of the idiosyncratic variance σ_θ^2 on these variables, given by the terms \hat{C}_0 and \hat{I}^k .

Stochastic discount factor. Consumption of investor i at date $t = 1$ is given by $C_{i,s} = R_{i,s}^a K_s / N$, where $R_{i,s}^a := \sum_j \omega_{i,j} R_{j,s}^a$. We can then write consumption of investor i as follows:

$$C_{i,s}(\sigma_\theta) = (1 + \psi_s(\sigma_\theta) \sigma_\theta \epsilon_{i,s}^p) \bar{R}_s^a(\sigma_\theta) \frac{K_s(\sigma_\theta)}{N}, \quad (\text{A.4})$$

where $\epsilon_{i,s}^p := \sum_j \omega_{i,j} \epsilon_{j,s}$, $\bar{R}_s^a(\sigma_\theta) = 1 - \delta + \alpha \Theta^\alpha K_s^{\alpha-1}(\sigma_\theta)$, and $\psi_s(\sigma_\theta) = \frac{\alpha \Theta^\alpha K_s^{\alpha-1}(\sigma_\theta)}{1 - \delta + \alpha \Theta^\alpha K_s^{\alpha-1}(\sigma_\theta)} \frac{1}{\Theta}$.

The second-order expansion of $C_{i,s}(\sigma_\theta)$ can be written as

$$C_{i,s}(\sigma_\theta) = C_s^* + C_s^* \psi_s^* \epsilon_{i,s}^p \sigma_\theta + \hat{C}_s \sigma_\theta^2 + o(\sigma_\theta^2), \quad (\text{A.5})$$

where $\psi_s^* = \psi_s(0)$, $\hat{C}_s = C_s^* (1 - (1 - \alpha) \Theta \psi_s^*) \frac{\hat{K}_s}{K_s^*}$, and $K_s(\sigma_\theta) = K_s^* + \hat{K}_s \sigma_\theta^2 + o(\sigma_\theta^2)$.

Taking a power k of $C_{i,s}$, we obtain

$$C_{i,s}^k(\sigma_\theta) = (C_s^*)^k + k(C_s^*)^k \psi_s^* \epsilon_{i,s}^p \sigma_\theta + \left[k(k-1)(C_s^*)^k (\psi_s^*)^2 \frac{(\epsilon_{i,s}^p)^2}{2} + k(C_s^*)^{k-1} \hat{C}_s \right] \sigma_\theta^2 + o(\sigma_\theta^2), \quad (\text{A.6})$$

The SDF for investor i can then be written as

$$M_{i,s} = \beta C_{i,s}^{-\gamma} C_0^\gamma = M_{i,s}^* - \gamma M_{i,s}^* \psi_s^* \epsilon_{i,s}^p \sigma_\theta + \hat{M}_{i,s} \sigma_\theta^2 + o(\sigma_\theta^2), \quad (\text{A.7})$$

where

$$\hat{M}_{i,s} = M_s^* \left[-\gamma \frac{\hat{C}_s}{C_s^*} + \frac{\gamma(\gamma+1)}{2} (\psi_s^*)^2 (\epsilon_{i,s}^p)^2 + \gamma \frac{\hat{C}_0}{C_0^*} \right] \quad (\text{A.8})$$

Firm's ROA. The ROA for firm j can be written as

$$R_{j,s}^a = R_s^{a,*} + R_s^{a,*} \psi_s^* \epsilon_{j,s} \sigma_\theta + \hat{R}_s^a \sigma_\theta^2 + o(\sigma_\theta^2), \quad (\text{A.9})$$

where $\hat{R}_s^a = -(1-\alpha)R_s^{a,*} \psi_s^* \Theta \frac{\hat{K}_s}{K_s^*}$.

A.3. Proof of Proposition 1

Proof. Take an arbitrary stockholder i (such that $j \in \mathcal{P}_i$) as the decision maker for that firm. This arbitrary choice is not a concern as the results on a firm's decisions will be shown are identity independent. The Euler equations that implicitly define the level of overall investment ($I = \sum_k I^k$) and its composition ($\chi = \frac{I^1}{I^0+I^1}$) can be written as follows:

$$\mathbb{E} \left[\beta C_{i,s}^{-\gamma} R_{j,s}^a \right] - C_0^{-\gamma} = 0, \quad \mathbb{E} \left[C_{i,s}^{-\gamma} R_{j,s}^a \varphi_s^e \right] = 0, \quad (\text{A.10})$$

where

$$C_{i,s} = (1 + \psi_s \sigma_\theta \epsilon_{i,s}^p) \bar{R}_s^a (1 + \chi \varphi_s^e) I, \quad C_0 = E_0 - I, \quad R_{j,s}^a = \bar{R}_s^a (1 + \psi_s \epsilon_{j,s} \sigma_\theta). \quad (\text{A.11})$$

Plugging the expression for consumption and return on assets into the Euler equations, we obtain

$$\mathbb{E} \left[\beta I^{-\gamma} (1 + \chi \varphi_s^e)^{-\gamma} \left(\bar{R}_s^a \right)^{1-\gamma} \eta_{i,s}^{-\gamma} \eta_{j,s} \right] - (E_0 - I)^{-\gamma} = 0 \quad (\text{A.12})$$

and

$$\mathbb{E} \left[\beta I^{-\gamma} (1 + \chi \varphi_s^e)^{-\gamma} (\bar{R}_s^a)^{1-\gamma} \eta_{i,s}^{-\gamma} \eta_{j,s} \varphi_s^e \right] = 0, \quad (\text{A.13})$$

where

$$\eta_{i,s} := 1 + \psi_s \epsilon_{i,s}^p \sigma_\theta \quad \text{and} \quad \eta_{j,s} := 1 + \psi_s \epsilon_{j,s} \sigma_\theta. \quad (\text{A.14})$$

Let us do a second-order asymptotic expansion around $\sigma_\theta^2 = 0$. We initially expand the following conditional expectation of interest,

$$\begin{aligned} \mathbb{E}_s \left[\beta I^{-\gamma} (1 + \chi \varphi_s^e)^{-\gamma} (\bar{R}_s^a)^{1-\gamma} (\eta_{i,s})^{-\gamma} \eta_{j,s} \right] &= \mu_s^* \left[1 - \gamma \frac{\hat{I}}{I^*} \sigma_\epsilon^2 - \gamma \frac{\hat{\chi}}{(1 + \chi \varphi_s^e)} \varphi_s^e \sigma_\theta^2 + (1 - \gamma) \frac{\hat{R}_s^a}{R_s^{a,*}} \sigma_\epsilon^2 \right. \\ &\quad \left. + \gamma(\gamma - 1) \frac{(\psi_s^*)^2}{2} \phi_u \sigma_\theta^2 \right], \end{aligned} \quad (\text{A.15})$$

where $\mu_s^* := \beta (I^*)^{-\gamma} (R_s^{a,*})^{1-\gamma} (1 + \chi^* \varphi_s^e)^{-\gamma}$ and we use the fact that $\text{Var}(\epsilon_{i,s}^p) = \text{Cov}(\epsilon_{i,s}^p, \epsilon_{j,s}) = \phi_u \sigma_\theta^2$.

Notice also that

$$\frac{\hat{R}_s^a}{R_s^{a,*}} = - (1 - \alpha) \psi_s^* \Theta \left[\frac{\hat{I}}{I^*} + \frac{\hat{\chi}}{(1 + \chi^* \varphi_s^e)} \varphi_s^e \right]. \quad (\text{A.16})$$

For the second term in Eq. (A.12), we have

$$(E_0 - I)^{-\gamma} = (E_0 - I^*)^{-\gamma} \left[1 + \gamma \frac{\hat{I}}{E_0 - I^*} \sigma_\theta^2 \right] + o(\sigma_\theta^2). \quad (\text{A.17})$$

Evaluating at $\sigma_\theta = 0$, we have $\mathbb{E} \left[\beta (I^*)^{-\gamma} (R_s^{a,*})^{1-\gamma} (1 + \chi^* \varphi_s^e)^{-\gamma} \right] = (E_0 - I^*)^{-\gamma}$, so $\mathbb{E} [\mu_s^*] = u'(C_0^*)$.

Combining these expressions, we obtain

$$\begin{aligned} \mathbb{E} \left[\mu_s^* \left(\gamma \left(\frac{\hat{I}}{I^*} + \frac{\hat{I}}{E_0 - I^*} \right) + (1 - \gamma) (1 - \alpha) \psi_s^* \Theta \left[\frac{\hat{I}}{I^*} + \frac{\hat{\chi}}{(1 + \chi^* \varphi_s^e)} \varphi_s^e \right] \right. \right. \\ \left. \left. + \gamma \frac{\hat{\chi}}{(1 + \chi^* \varphi_s^e)} \varphi_s^e - \gamma(\gamma - 1) \frac{(\psi_s^*)^2}{2} \phi_u \right) \right] = 0. \end{aligned} \quad (\text{A.18})$$

Now, let's focus on Eq. (A.13) and expand the following conditional expectation

$$\mathbb{E}_s \left[\beta (I^*)^{-\gamma} (\bar{R}_s^a)^{1-\gamma} (1 + \chi \varphi_s^e)^{-\gamma} \eta_{i,s}^{-\gamma} \eta_{j,s} \varphi_s^e \right] = \mu_s^* \varphi_s^e \left[1 + (1 - \gamma) \frac{\hat{R}_s^a}{\bar{R}_s^{a,*}} \sigma_\theta^2 - \gamma \frac{\hat{\chi}}{(1 + \chi \varphi_s^e)} \varphi_s^e \sigma_\theta^2 + \gamma (\gamma - 1) \frac{(\psi_s^*)^2}{2} \phi_u \sigma_\theta^2 \right]. \quad (\text{A.19})$$

Notice that, evaluating at $\sigma_\theta = 0$, $\mathbb{E} [\mu_s^* \varphi_s^e] = 0$. Therefore, any state-invariant term in Eq. (A.19) cancels out once averaged out across states. The ensuing equation can also be written in a covariance form.

Organizing the system formed by Eq. (A.18) and the unconditional expectation of Eq. (A.19) in matrix form, we obtain

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} \hat{I} \\ \hat{\chi} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad (\text{A.20})$$

where the coefficients $a_{i,j}$ and intercepts b_i are provided as follows. First,

$$a_{1,1} = \mathbb{E} \left[\mu_s^* \left(\gamma \left(\frac{1}{E - I^*} + (1 - (1 - \alpha) \psi_s^* \Theta) \frac{1}{I^*} \right) + (1 - \alpha) \frac{\psi_s^* \Theta}{I^*} \right) \right] > 0. \quad (\text{A.21})$$

Additionally,

$$a_{1,2} = \mathbb{E} \left[\mu_s^* \left((1 - \alpha) \psi_s^* \Theta + \gamma (1 - (1 - \alpha) \psi_s^* \Theta) \frac{\varphi_s^e}{(1 + \chi^* \varphi_s^e)} \right) \right] < 0, \quad (\text{A.22})$$

in which the sign follows after some manipulation,¹⁵ and

$$a_{2,1} = (1 - \gamma) (1 - \alpha) \text{Cov} \left(\mu_s^* \varphi_s^e, \frac{\psi_s^* \Theta}{I^*} \right). \quad (\text{A.23})$$

For characterizing the sign of $a_{2,1}$, first notice that, $\mu_H^* < \mu_L^*$, $\varphi_H^e > 0 > \varphi_L^e$, and $\psi_H < \psi_L$ (countercyclical idiosyncratic risk). So, $\text{Cov} \left(\mu_s^* \varphi_s^e, \frac{\psi_s^* \Theta}{I^*} \right) < 0$ and it follows that $a_{2,1} > 0$ whenever $\gamma > 1$.

We also obtain the following:

$$a_{2,2} = \mathbb{E} \left[\mu_s^* \frac{(\varphi_s^e)^2}{(1 + \chi \varphi_s^e)} \left((1 - \alpha) \psi_s^* \Theta + \gamma (1 - (1 - \alpha) \psi_s^* \Theta) \right) \right] > 0. \quad (\text{A.24})$$

¹⁵Note that $a_{1,2} = \text{Cov} \left(\mu_s^* \varphi_s^e, \frac{\gamma(1-\delta) + (\alpha\gamma + (1-\alpha))\alpha\Theta^\alpha K_s^{\alpha-1}}{(1-\delta)K_s + \alpha\Theta^\alpha K_s^\alpha} \right) I^* < 0$.

Last, $b_1 = \frac{\gamma(\gamma-1)}{2}\phi_u\mathbb{E}\left[\mu_s^*(\psi_s^*)^2\right]$ and $b_2 = \frac{\gamma(\gamma-1)}{2}\phi_u\text{Cov}\left(\mu_s^*\varphi_s^e, (\psi_s^*)^2\right)$. Whenever $\gamma > 1$, $b_1 > 0$ and $b_2 < 0$. Let A be the matrix defined by $a_{i,j}$ above. Then, $\gamma > 1$ ensures that $\det(A) > 0$ and

$$\hat{I} = \frac{a_{2,2}b_1 - a_{1,2}b_2}{\det(A)}, \quad (\text{A.25})$$

which has an ambiguous sign. Notice that if $\hat{\chi} = 0$ as a constraint or $b_2 \rightarrow 0$ (which occurs when $\delta \rightarrow 1$ and ψ_s^* converges to a constant, becoming acyclical), then $\hat{I} > 0$.

For $\hat{\chi}$, we have

$$\hat{\chi} = \frac{a_{1,1}b_2 - a_{2,1}b_1}{\det(A)} < 0. \quad (\text{A.26})$$

□

Appendix B. Appendix for Section 3

B.1. Proof of Propositions 2 and 3

Proof. We first consider the welfare implications of a general perturbation (κ_0, κ_1) . Then we study two specific cases: in one case, the perturbation affects the aggregate level of investment; in another, it only alters the aggregate risk taking but leaves the total investment unchanged.

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

$$\begin{aligned} V'(\Delta) &= \sum_k \left(-u'(C_0) + \beta \mathbb{E} \left[u'(C_{i,s}) \left(R_{i,s}^a(\Delta) \varphi_s^k \right) \right] \right) \kappa_k \\ &\quad + \beta \sum_k \left(\mathbb{E} \left[u'(C_{i,s}) \left(\left(\frac{\partial \bar{R}_s^a(\Delta)}{\partial K_s} K_s(\Delta) \frac{\theta_{i,s}}{\Theta} + (1-\alpha) \alpha \Theta^\alpha K_s^{\alpha-1}(\Delta) \right) \right) \varphi_s^k \right] \right) \kappa_k. \end{aligned} \quad (\text{B.1})$$

Notice that $\frac{\partial \bar{R}_s^a(\Delta)}{\partial K_s} K_s(\Delta) = -(1-\alpha) \alpha \Theta^\alpha K_s^{\alpha-1}(\Delta)$. When computed under the laissez-faire allocation, the first term in Eq. (B.1) vanishes and

$$\frac{\partial \bar{R}_s^a(0)}{\partial K_s} K_s(0) = -(1-\alpha) \psi_s \Theta \bar{R}_s(0), \quad (\text{B.2})$$

allowing us to obtain

$$V'(0) = -(1-\alpha)\beta \sum_k \left(\mathbb{E} \left[u'(C_{i,s}) \psi_s \epsilon_{i,s}^p \sigma_\theta \bar{R}_s^a(0) \varphi_s^k \right] \right) \kappa_k \quad (\text{B.3})$$

$$= -(1-\alpha)\beta \sum_k \left(\mathbb{E} \left[u'(C_{i,s}) (\eta_{i,s} - 1) \bar{R}_s^a(0) \varphi_s^k \right] \right) \kappa_k, \quad (\text{B.4})$$

where $\eta_{i,s}$ is defined as in Eq. (A.14).

Therefore, for the general perturbation,

$$\frac{V'(0)}{u'(C_0)} = -(1-\alpha) \mathbb{E} \left[\text{Cov}_s(M_{i,s}, R_{i,s}^a) (\kappa_0 + \kappa_1 \varphi_s^1) \right], \quad (\text{B.5})$$

using the fact that $\mathbb{E}_s[R_{i,s}^a - \bar{R}_s^a] = \mathbb{E}_s[(\eta_{i,s} - 1)\bar{R}_s^a] = 0$.

The covariance above can be written as

$$\begin{aligned} \text{Cov}_s(M_{i,s}, R_{i,s}^a) &= \text{Cov}_s \left(-\gamma M_{i,s}^* \psi_s^* \epsilon_{i,s}^p \sigma_\theta, R_{i,s}^{a,*} \psi_s^* \epsilon_{i,s}^p \sigma_\theta \right) + o(\sigma_\theta^2) \\ &= -\gamma \phi_u \sigma_s^2 M_{i,s}^* \bar{R}_s^{a,*} + o(\sigma_\theta^2). \end{aligned} \quad (\text{B.6})$$

The derivative of the value function can then be written as

$$\frac{V'(0)}{u'(C_0)} = (1-\alpha)\gamma\phi_u \mathbb{E} \left[\beta \frac{u'(C_s^*)}{u'(C_0^*)} R_{i,s}^{a,*} \sigma_s^2 (\kappa_0 + \kappa_1 \varphi_s^1) \right] + o(\sigma_\theta^2). \quad (\text{B.7})$$

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

$$\frac{V'(0)}{u'(C_0)} = (1-\alpha)\gamma\phi_u \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \sigma_s^2 (\kappa_0 + \kappa_1 \varphi_s^1) \right] + o(\sigma_\theta^2). \quad (\text{B.8})$$

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

$$\mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \right] = 1. \quad (\text{B.9})$$

Finally, define the risk-neutral probabilities as follows¹⁶

$$\mathbb{E}^{\mathbb{Q}} [X_s] \equiv \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a X_s \right], \quad (\text{B.10})$$

for any random variable X_s .

This allows us to write

$$\frac{V'(0)}{u'(C_0)} = (1 - \alpha) \gamma \phi_u \mathbb{E}^{\mathbb{Q}} \left[\sigma_s^2 (\kappa_0 + \kappa_1 \varphi_s^1) \right] + o(\sigma_\theta^2). \quad (\text{B.11})$$

Underinvestment. Take $\kappa_0 = 1$ and $\kappa_1 = 0$. Then, for the first part of the proposition, notice that Eq. (B.5) implies that

$$\frac{V'(0)}{u'(C_0)} = - (1 - \alpha) \mathbb{E} \left[\beta \frac{\bar{C}_s^{-\gamma}}{C_0^{-\gamma}} \bar{R}_s \text{Cov}_s \left(\eta_{i,s}^{-\gamma}, \eta_{i,s} \right) \right] > 0, \quad (\text{B.12})$$

because the conditional covariance above involves the random variable $\eta_{i,s}$ and a strictly decreasing function of itself. For the asymptotic approximation, notice that $\frac{V'(0)}{u'(C_0)} = (1 - \alpha) \gamma \phi_u \mathbb{E}^{\mathbb{Q}} [\sigma_s^2]$. Corollary 1 follows immediately by imposing $\alpha = 1$ in the expression above.

Last, we show that the idiosyncratic variance risk premium is positive:

$$\mathbb{E}^{\mathbb{Q}} [\sigma_s^2] - \mathbb{E} [\sigma_s^2] = \mathbb{E} [M_{i,s} R_{i,s}^a \sigma_s^2] - \mathbb{E} [M_{i,s} R_{i,s}^a] \mathbb{E} [\sigma_s^2] \quad (\text{B.13})$$

$$= \text{Cov} \left(\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a, \sigma_s^2 \right) > 0, \quad (\text{B.14})$$

where the inequality follows from $u'(C_{i,s})$, $R_{i,s}^a$, and σ_s^2 being all countercyclical.

Excessive aggregate risk-taking. Taking $\kappa_0 = \frac{1}{\sqrt{\text{Var}^{\mathbb{Q}}[\varphi^1]}}$ and $\kappa_1 = -\frac{1}{\sqrt{\text{Var}^{\mathbb{Q}}[\varphi^1]}}$, we have that Eq. (B.5) implies that

$$\frac{V'(0)}{u'(C_0)} = (1 - \alpha) \mathbb{E} \left[\beta \frac{\bar{C}_s^{-\gamma}}{C_0^{-\gamma}} \bar{R}_s \text{Cov}_s \left(\eta_{i,s}^{-\gamma}, \eta_{i,s} \right) \varphi_s^e \kappa_0 \right]. \quad (\text{B.15})$$

¹⁶Note that $\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a$ 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.

Then, notice that the Euler equation for the composition of investment implies

$$\mathbb{E} \left[\beta \frac{\bar{C}_s^{-\gamma}}{C_0^{-\gamma}} \bar{R}_s \varphi_s^e \mathbb{E}_s \left[\eta_{i,s}^{-\gamma} \eta_{i,s} \right] \right] = 0. \quad (\text{B.16})$$

Then, using that $\mathbb{E}_s [\eta_{i,s}] = 1$ and the last two equations, we obtain

$$\frac{V'(0)}{u'(C_0)} = -(1-\alpha) \mathbb{E} \left[\beta \frac{\bar{C}_s^{-\gamma}}{C_0^{-\gamma}} \bar{R}_s \mathbb{E}_s \left[\eta_{i,s}^{-\gamma} \right] \varphi_s^e \right] \kappa_0 > 0, \quad (\text{B.17})$$

because $\beta \frac{\bar{C}_l^{-\gamma}}{C_0^{-\gamma}} > \beta \frac{\bar{C}_l^{-\gamma}}{C_0^{-\gamma}}$, $\varphi_h^e > 0 > \varphi_l^e$, $\eta_{i,l}$ is a mean-preserving spread of $\eta_{i,h}$ and the conditional expectation term above involves a convex function.

For the asymptotic approximation, notice that

$$\begin{aligned} \frac{V'(0)}{u'(C_0)} &= -(1-\alpha) \gamma \phi_u \frac{\mathbb{E}^{\mathbb{Q}} [\sigma_s^2 (\varphi_s^1 - \mathbb{E}^{\mathbb{Q}}[\varphi^1])] }{\sqrt{\text{Var}^{\mathbb{Q}}[\varphi^1]}} \\ &= -(1-\alpha) \gamma \phi_u \frac{\text{Cov}^{\mathbb{Q}} (\sigma_s^2, \varphi_s^1)}{\sqrt{\text{Var}^{\mathbb{Q}}[\varphi^1]}}, \end{aligned} \quad (\text{B.18})$$

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

$$\frac{V'(0)}{u'(C_0)} = -(1-\alpha) \gamma \phi_u \frac{1}{\sqrt{q_h q_l} (\varphi_h^1 - \varphi_l^1)} \left[q_h q_l \sigma_h^2 (\varphi_h^1 - \varphi_l^1) - q_h q_l \sigma_l^2 (\varphi_h^1 - \varphi_l^1) \right] \quad (\text{B.19})$$

$$= (1-\alpha) \gamma \phi_u \sqrt{q_h q_l} (\sigma_l^2 - \sigma_h^2), \quad (\text{B.20})$$

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

The idiosyncratic variance risk premium can be written as

$$\mathbb{E}^{\mathbb{Q}}[\sigma_s^2] - \mathbb{E}[\sigma_s^2] = (q_l - p_l)(\sigma_l^2 - \sigma_h^2) > 0, \quad (\text{B.21})$$

given $q_l > p_l$ and $\sigma_l > \sigma_h$.

Combining the previous two equations, we obtain the last expression in the proposition. \square

B.2. Extensions

See the online supplement, Section OS.1.

Appendix C. Appendix for Section 4

In Section 4, we perform a quantification of the welfare impact of the idiosyncratic risk externalities. In this appendix, we provide additional detail.

C.1. Estimating the idiosyncratic risk premium

From Eq. (29), 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:

$$r_{i,t+1}^e = \lambda_0 + \lambda_{id}\mathbb{E}_t[\sigma_{i,t+1}^2] + \boldsymbol{\lambda}' \mathbf{X}_{i,t} + \epsilon_{i,t+1}, \quad i = 1, \dots, N, \quad t = 1, \dots, T - 1, \quad (\text{C.1})$$

where $r_{i,t+1}^e$ is the realized excess return on stock i in period $t + 1$, $\mathbb{E}_t[\sigma_{i,t+1}^2]$ is the expected variance of the idiosyncratic return in $t + 1$ conditional on information in t , and $\mathbf{X}_{i,t}$ 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 $\mathbb{E}_t[\sigma_{i,t+1}^2]$, to which we refer 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.

Ang et al. (2006) find 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 variance. Fu (2009) points out the importance of accounting for mean-reversion in volatility, as the results in Ang et al. (2006) may be biased in this case. See also Mehra et al. (2019), Spiegel and Wang (2007) and Eiling (2013), that obtain positive premium estimates with related approaches.

Sample and variables. Following the convention, we test specification (C.1) 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 the 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

Table C.1

Summary statistics. This table provides the summary of 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 from 1963M07 through 2018M12. The selected characteristics include: $\mathbb{E}_{t-1}[\sigma_{i,t}^2]$, 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; $R_{t-7 \rightarrow t-2}$, 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.

Characteristics	Mean	S.D.	Percentiles				
			10th	25th	50th	75th	90th
$r_{i,t}$ (%)	1.060	14.010	-13.905	-6.015	0.000	6.977	16.129
$\mathbb{E}_{t-1}[\sigma_{i,t}^2]$ (%)	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
$R_{t-7 \rightarrow t-2}$	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

return as the residual excess return that is unexplained by Fama and French's (1993) three factors. Specifically, we postulate the following representation of excess returns:

$$r_{i,t}^e = \alpha_i + \beta_{i,mkt} r_{m,t}^e + \beta_{i,smb} SMB_t + \beta_{i,hml} HML_t + \varepsilon_{i,t}, \text{ where } \varepsilon_{i,t} \sim N(0, \hat{\sigma}_{i,t}^2) \quad (\text{C.2})$$

$$\ln \hat{\sigma}_{i,t}^2 = a_i + \sum_{j=1}^p b_{i,j} \ln \hat{\sigma}_{i,t-j}^2 + \sum_{k=1}^q c_{i,k} \left\{ \theta \left(\frac{\varepsilon_{i,t-k}}{\hat{\sigma}_{i,t-k}} \right) + \nu \left[\left| \frac{\varepsilon_{i,t-k}}{\hat{\sigma}_{i,t-k}} \right| - \sqrt{\frac{2}{\pi}} \right] \right\},$$

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 C.1, the median expected idiosyncratic variance in our sample is 0.944%, similar to that reported in Fu (2009). In Fig. 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; ME , the market capitalization of the issuing firm; BM , the book-to-market ratio of the issuing firm; $R_{t-7 \rightarrow t-2}$, 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.

Specifically, 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. 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. 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 year-end information from the Compustat database. 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. 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 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%.

Fama-MacBeth regressions. Within this sample of stocks, we estimate Eq. (C.1) 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 C.2.

We start by replicating some well-documented results in the literature. The results we report in column 1 of Table C.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 C.2, we

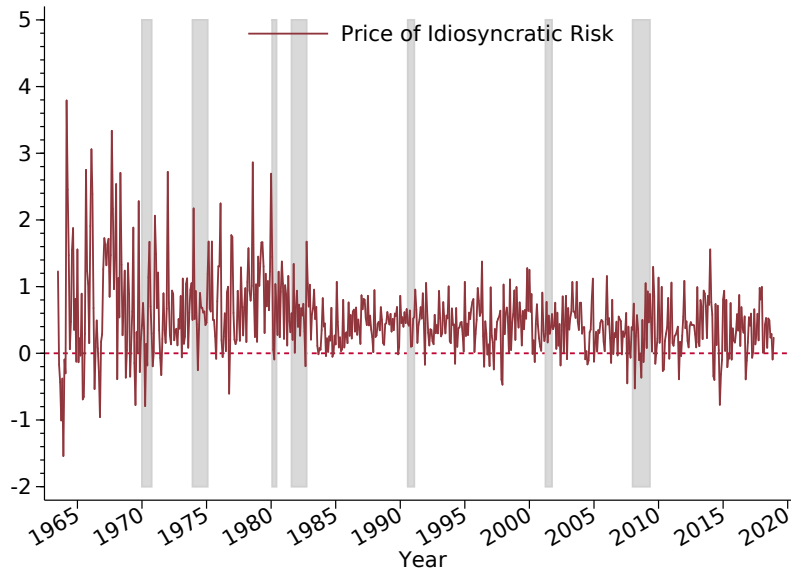
Table C.2

Fama-MacBeth regressions. This table reports the estimation results of Fama-MacBeth regressions specified as $r_{i,t+1}^e = a + \lambda_{ivar} \mathbb{E}_t[\sigma_{i,t+1}^2] + \boldsymbol{\lambda}' \mathbf{X}_{i,t} + \epsilon_{i,t+1}$, where $r_{i,t+1}^e (\equiv r_{i,t+1} - r_{f,t})$ is the return on stock i in excess of the one-month Treasury bill rate in month $t + 1$, and $\mathbb{E}_t[\sigma_{i,t+1}^2]$ 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; $R_{t-6 \rightarrow t-1}$, 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.

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

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 $\mathbb{E}_t[\sigma_{i,t+1}^2]$ 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 report regressions that include other characteristics to control for exposure to common risk factors. We find that the explanatory power of idiosyncratic risk becomes even stronger: the slopes for $\mathbb{E}_t[\sigma_{i,t+1}^2]$ are always more than 10 standard errors away 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.

Fig. C.1. 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 the NBER.



Stability of the price of risk estimation. We plot, in Fig. C.1, 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 exhibited 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. Return variance decomposition

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

$$\begin{aligned} \text{Var}_t[r_{j,t+1}] &= \text{Var}_t[r_{m,t+1}] + \text{Var}_t[v_{j,t+1}] + 2\text{Cov}_t(r_{m,t}, v_{j,t}) \\ &= \text{Var}_t[r_{m,t+1}] + \text{Var}_t[v_{j,t+1}] + 2(\beta_{j,t} - 1)\text{Var}_t(r_{m,t+1}), \end{aligned} \quad (\text{C.3})$$

Let $\bar{\sigma}_t^2 \equiv \sum_j \omega_{m,j} \text{Var}_t[r_{j,t+1}]$ and $\bar{\sigma}_{id,t}^2 \equiv \sum_j \omega_{m,j} \text{Var}_t[v_{j,t+1}]$, then

$$\bar{\sigma}_t^2 = \sigma_{m,t}^2 + \bar{\sigma}_{id,t}^2, \quad (\text{C.4})$$

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

Appendix D. Appendix for Section 5

D.1. Modified Setup

We assume that the planner controls investment and risk-taking through financial regulation. We introduce a set 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 particular locations. We assume that each intermediary $j \in \mathcal{J} = \{\frac{1}{N}, \frac{2}{N}, \dots, 1\}$ is in a bilateral relationship with firm j and that the terms of the lending contract are determined through bargaining. For the sake of simplicity, we assume that the intermediary has all the bargaining power, so firms make no profits in equilibrium. Given full rent-extraction by intermediaries, we simply assume that non-financial firms are entirely bank-financed and that the investors then choose a portfolio of financial firms within their limited-participation sets. The assumption that the intermediary has all the bargaining power simplifies the exposition, but it is not essential for the argument.

Financial intermediaries' problem. Each intermediary $j \in \mathcal{J}$ 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 intermediary's borrowing costs. Formally, the intermediary solves the problem

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

subject to

$$D_j \leq (1 - \delta) \sum_k I_j^k \phi_L^k, \quad \sum_k I_j^k - P_d D_j \geq \sum_k \omega^k I_j^k. \quad (\text{D.2})$$

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

$$\max_{C_{i,0}, \{\omega_{i,j}\}_{j \in \mathcal{J}}} u(C_{i,0}) + \beta \mathbb{E} [u(C_{i,s})], \quad (\text{D.3})$$

subject to non-negativity conditions on consumption, the participation constraint, $\omega_{i,j} = 0, \forall j \notin \mathcal{P}_i$, the summation constraint on portfolio weights, $\sum_{j \in \mathcal{J}} \omega_{i,j} = 1$, and budget constraint,

$$C_{i,s} = R_{i,s}(E_0 - T - C_{i,0}) + T_{w,s}, \quad (\text{D.4})$$

where

$$R_{i,s} \equiv \Psi_i \sum_{j \in \mathcal{J}} \omega_{i,j} \frac{R_{j,s}^a K_{j,s} - D_j}{P_{e,j}} + (1 - \Psi_i) \frac{1}{P_d}, \quad (\text{D.5})$$

is the return on the investor's portfolio, Ψ_i is the portfolio weight on risky assets, $\omega_{i,j}$ is the share of investor i 's equity investments allocated on intermediary j 's stock, T is a lump-sum levy used to finance the debt tax benefit, and $T_{w,s}$ is a lump-sum transfer from workers.

The equilibrium definition is provided in the Online Appendix. Section OS.5 of that supplement offers an implementation result, demonstrating that a planner can use the tax benefit, 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 benefit and a

risk-weighted capital requirement constraint is imposed on intermediaries.

D.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 Supplement OS.5 to characterize how a tax benefit on debt 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. The magnitude of these external effects are related to the optimal level of the policy instruments.

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.

We write the relaxed planning program as

$$\max_{I, \chi, \{\omega_{i,j}\}_{j \in \mathcal{J}}, \{T_{w,s}\}_s} u(E_0 - I) + \beta \mathbb{E} \left[u \left(\sum_j \omega_{i,j} R_{j,s}^a K_{j,s} + \frac{T_{w,s}}{N} \right) \right], \quad (\text{D.6})$$

subject to the participation constraint, $\omega_{i,j} = 0, \forall j \notin \mathcal{P}_i$, the constraint on the sum of portfolio weights, $\sum_{j \in \mathcal{J}} \omega_{i,j} = 1$, and

$$E \left[u_w \left((1 - \alpha) (\Theta K_s)^\alpha - T_{w,s} \right) \right] \geq \underline{u}^w, \quad (\text{D.7})$$

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

In this planning problem, all constraints on feasibility and the distribution of consumption across agents are taken into account, including the same limited participation in idiosyncratic risk-sharing as before. Additionally, Constraint (D.7) 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 $T_{w,s}$, 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 5 (Optimal Policy). *The necessary first-order conditions of Problem (D.6) can be summarized as:*

i. *A planner's investment Euler equation,*

$$1 = \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{j,s}^a \frac{K_{j,s}}{I(1 - IRE_I)} \right], \quad (\text{D.8})$$

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

ii. *A planner's risky technology Euler equation,*

$$\mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \varphi_s^e \right] = IRE_\chi, \quad (\text{D.9})$$

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

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

$$\mathbb{E} \left[u'(C_{i,s}) R_{j,s}^a K_s \right] = \mathbb{E} \left[u'(C_{i,s}) R_{i,s}^a K_s \right]. \quad (\text{D.10})$$

Proof. At the end of this section.

Proposition 5 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 2 and 3. The third condition characterizes the planner's optimal portfolio. It coincides with the condition for private investors' portfolios that establishes the optimality of equally weighted portfolios. 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 benefit and the risk weights. Comparing the investment Euler equation for the financial intermediary with the corresponding one for the

planner, we obtain

$$P_d d\tau_d = IRE_I, \quad d \equiv \frac{D}{I}. \quad (\text{D.11})$$

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

$$(\omega^1 - \omega^0)\tau_d P_d = IRE_\chi. \quad (\text{D.12})$$

Proof of Proposition 5. Consider Program (D.6). The first-order condition with respect to $T_{w,s}$ is

$$\beta \mathbb{E}_s [u'(C_{i,s})] = N\mu_w u'_w(C_{w,s}). \quad (\text{D.13})$$

The optimality condition for the portfolio allocation is

$$\mathbb{E} [u'(C_{i,s}) R_{j,s}^a K_s] = \mathbb{E} [u'(C_{i,s}) R_{i,s}^a K_s]. \quad (\text{D.14})$$

The first-order condition for I is

$$-u'(C_0) + \beta \mathbb{E} \left[u'(C_{i,s}) R_{i,s}^a \frac{K_{j,s}}{I} \right] + \beta \mathbb{E} \left[u'(C_{i,s}) \frac{\partial R_{i,s}^a}{\partial I} K_{j,s} \right] + N\mu_w \mathbb{E} \left[u'_w(C_{w,s}) \alpha (1 - \alpha) \Theta^\alpha K_s^{\alpha-1} \frac{K_s}{I} \right] = 0, \quad (\text{D.15})$$

where $\frac{\partial R_{i,s}^a}{\partial I} = -(1 - \alpha)\alpha\bar{\theta}_i\Theta^{\alpha-1}K_s^{\alpha-2}\frac{K_s}{I}$.

Using the optimality condition for $T_{w,s}$, we obtain

$$1 = \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \frac{K_{j,s}}{I} \right] - (1 - \alpha) \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} (R_{i,s}^a - \bar{R}_s) \frac{K_{j,s}}{I} \right] \quad (\text{D.16})$$

$$= \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \frac{K_{j,s}}{I(1 - IRE_I)} \right], \quad (\text{D.17})$$

given that the last term in the first equation is

$$IRE_I \equiv -(1 - \alpha) \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} (R_{i,s}^a - \bar{R}_s) \frac{K_{j,s}}{I} \right]. \quad (\text{D.18})$$

From the optimality condition for the portfolio allocation, we can replace $R_{i,s}^a$ by $R_{j,s}^a$ for any

$j \in \mathcal{P}_i$, in Equation (D.17), delivering Equation (D.8) from the proposition's statement.

Last, the optimality condition for the share invested in the risky technology is

$$\beta \mathbb{E} \left[u'(C_{i,s}) R_{i,s}^a \varphi_s^e \right] + N \mu_w \mathbb{E} \left[u'_w(C_{w,s}) \alpha (1 - \alpha) \Theta^\alpha K_s^{\alpha-1} \varphi_s^e \right] + \beta \mathbb{E} \left[u'(C_{i,s}) \frac{\partial R_{i,s}^a}{\partial \chi} \frac{K_{j,s}}{I} \right] = 0, \quad (\text{D.19})$$

where $\frac{\partial R_{i,s}^a}{\partial \chi} = -(1 - \alpha) \alpha \bar{\theta}_i \Theta^{\alpha-1} K_s^{\alpha-2} \varphi_s^e I$.

Using the optimality condition for $T_{w,s}$, we obtain

$$\mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} R_{i,s}^a \varphi_s^e \right] = IRE_\chi, \quad (\text{D.20})$$

where

$$IRE_\chi \equiv (1 - \alpha) \mathbb{E} \left[\beta \frac{u'(C_{i,s})}{u'(C_0)} \left(R_{i,s}^a - \bar{R}_s^a \right) \varphi_s^e \right]. \quad (\text{D.21})$$

The approximations for IRE_I and IRE_χ are the same as in Section 3. □