Endogenous Dispersion and Volatility in Stock Returns*

Wukuang Cun†
Shanghai University of Finance and Economics

Junjie Xia‡
Central University of Finance and Economics and Peking University

February 25, 2026

Abstract

Measures of stock market volatility, such as the cross-sectional dispersion of firm-level stock returns and the time-series volatility of market returns, fluctuate countercyclically. While this phenomenon is often used as evidence that exogenous rises in uncertainty lead to economic recessions, we show that first-moment productivity shocks can cause countercyclical fluctuations in stock market volatility. We study a dynamic model with costly state verification, allowing for ex-ante monitoring that can credibly compel borrowers to reveal true investment outcomes to lenders, thereby reducing information asymmetry. Interpreting these financial contracts as external equities, we show that the calibrated model, with only first-moment shocks, can generate cyclical fluctuations in equity return volatility and financial conditions, with correlation coefficients between pairs of these measures quantitatively in line with the data.

Keywords: Asymmetric information, costly state verification, stock market volatility, business cycles

*We would like to thank Roberto Chang, Todd Keister, Vincenzo Quadrini, Joel David, Shenzhe Jiang and Qi Sun for their insightful comments and suggestions. Junjie Xia thanks support from the National Natural Science Foundation of China (72342031; 72473172) and the Major Program of the National Social Science Foundation of China (24ZDA057). All remaining errors are ours.

†Wukuang Cun is affiliated with the Shanghai University of Finance and Economics, China. Email: cunwukuang@mail.shufe.edu.cn

‡Junjie Xia is affiliated with the Central University of Finance and Economics and Peking University, China. Email: junjiexia@nsd.pku.edu.cn

1 Introduction

Measures of stock market volatility, such as the cross-sectional dispersion of individual stock returns and the time-series volatility of market returns, tend to rise during economic downturns (Campbell and Lettau (1999)). The significant countercyclical movements of these measures are often used as evidence supporting the notion that uncertainty shocks – exogenous increases in uncertainty – negatively impact aggregate economic activities and lead to recessions (see, for example, Bloom (2009), Gilchrist et al. (2014) and Basu and Bundick (2017), among many others).[1]

However, there is no consensus on whether higher uncertainty is an exogenous source of recessions or an endogenous response to economic fundamentals. In contrast with the studies mentioned above, there is also a strand of literature that view higher uncertainty (volatility) as an outcome, rather than a cause, of economic recessions. Examples of models based on information and financial frictions include Van Nieuwerburgh and Veldkamp (2006), Bachmann and Moscarini (2011), Gorton and Ordoñez (2014) and Ai et al. (2020). Further, empirical studies, such as Ludvigson et al. (2021) and Cesa-Bianchi et al. (2020), found that the observed countercyclical fluctuations of uncertainty measures can be due to first-moment shocks to economic fundamentals. In particular, as pointed out by Jurado et al. (2015), stock market volatility and cross-sectional dispersion in individual stock returns may change over the business cycle even if there is no change in uncertainty about economic fundamentals.

We propose a mechanism through which a negative first-moment shock to the aggregate productivity leads to increases in both the cross-sectional dispersion of individual stock returns and the time series volatility of market returns. Our study complements the literature by providing an alternative explanation for the observed countercyclical movements of measures of stock market volatility.

We consider a dynamic model incorporated with the costly state verifi-

[1] Bloom (2009) estimated a vector autoregression (VAR) model and found that a shock to the market volatility index leads to a rapid fall of production and employment. Gilchrist et al. (2014) conducted a VAR analysis showing that a shock to stock return volatility results in a widening credit spread and decreases in both output and investment. Using a structural VAR model, Basu and Bundick (2017) identified an uncertainty shock as an exogenous increase in the implied volatility of future stock returns.

cation (CSV henceforth) of Townsend (1979).[2] However, we deviate from conventional CSV models by allowing for both *ex-ante monitoring* and *ex-post monitoring*. Ex-ante monitoring refers to setting up a mechanism in advance to credibly force borrowers to reveal true investment outcomes to lenders.[3] In contrast, ex-post monitoring happens only when the promised payment is not made and thus the project is liquidated by lenders, as in conventional CSV models. Borrowers can utilize combinations of the two monitoring technologies. Although ex-ante monitoring is also costly, it enables payment to lenders to be made contingent on revealed information, thereby reducing the likelihood of costly liquidation (ex-post monitoring).

Another important feature of our model is that we allow borrowers to accumulate net worth, and thus the financial conditions of borrowers play an important role in financial contracting. When borrowers' net worth declines after adverse (first-moment) shocks, the costs of external financing increase. Consequently, optimal contracts call for more intensive use of ex-ante monitoring, requiring payments to lenders be contingent on revealed information. This mitigates information friction and reduces the likelihood of costly liquidation (ex-post monitoring). However, as a byproduct, the optimal financial contracts become more 'equity-like', with lenders assuming more aggregate and idiosyncratic risks, indicating higher volatility in lenders' returns from the contracts.

It is worth noting that this mechanism differs from risk-sharing, since it is not about smoothing consumption between borrowers and lenders. In fact, in the model both parties are assumed to be risk-neutral. However, it is optimal to distribute more risks to lenders for a borrower with lower net worth, since it helps reduce the likelihood of costly liquidation (ex-post monitoring). Thus, in our model, it is the minimization of monitoring costs, rather than risk-sharing, that determines the distribution of risk between lenders and borrowers.

[2] Krasa and Villamil (2000) show that the CSV model can be viewed as a reduced form of a costly enforcement model in which enforcement is chosen optimally by investors as part of a perfect Bayesian Nash equilibrium.

[3] The ex-ante monitoring in our model is similar to the observable return investment technology in Chang (1999), Seward (1990) and Boyd and Smith (1998, 1999). However, our paper differs from theirs in some important ways. While their papers study optimal investment and capital structure in static environments, we consider a dynamic model in which borrowers are allowed to accumulate net worth over time, and investigate how information asymmetry and optimal financial contracts vary with borrowers' net worth over business cycles.

The financial contracts in our model are interpreted as external equities, while the borrowers correspond to entrepreneur-managers, who not only operate firms but also hold internal equities of the firms. As pointed out by Leland and Pyle (1977) and Myers and Majluf (1984), issuing external equity can be costly due to information asymmetry between the manager and the potential buyers of external equity.[4] Our model explores the endogenous nature of corporate information friction associated with external financing by allowing for endogenous variation in the intensity of ex-ante monitoring.

Our model predicts that both the cross-sectional dispersion of individual equity returns and the volatility of market return increase following (first-moment) negative shocks to the aggregate productivity. This is due to the following reasons:

First, as mentioned before, as borrowers' net worth (i.e., internal equity of entrepreneur-managers) declines after adverse first-moment shocks, external financing becomes more difficult. Thus, ex-ante monitoring is used more intensively to facilitate external equity financing, requiring payments be contingent on revealed information. As a byproduct, more aggregate and idiosyncratic risks are distributed to lenders (i.e., investors of external equities), indicating higher volatility in equity returns. This result is consistent with the long-standing notion that individual stock return volatility is closely related to the flow of firm-specific information (Campbell et al. (2023)).

Second, in the model, following adverse first-moment aggregate shocks, heterogeneity in borrowers' financing decisions increases, which also contributes to the increased cross-sectional dispersion of equity returns. This is due to the fact that non-financially constrained borrowers always make the same financing decisions regardless of their wealth levels, while financially constrained borrowers make different decisions according to their wealth levels. After an adverse shock, borrowers' wealth distribution is shifted to the left. This increases the fraction of financially constrained borrowers, leading to greater heterogeneity in their decisions and thus larger cross-sectional dispersion in equity returns.

The calibrated model, subject to only first-moment productivity shocks,

[4] To reduce information asymmetry and thus the costs of external financing, firms tend to increase information disclosure when anticipating equity offerings (see Section 5.2.1 in Healy and Palepu (2001)).

can generate cyclical patterns of stock return volatility and financial conditions that are quantitatively inline with the data. Our exercises focus on three measures, namely the cross-sectional standard deviation of individual stock returns, the volatility of market return, and the tightness of the economy-wide financial conditions. The model generated correlation coefficients between pairs of these measures align quantitatively with their empirical counterparts. Further, our impulse response exercise shows that following a first-moment shock with no persistence, the increases in stock return volatility are not only significant but also persistent. This persistence arises because stock return volatility is largely determined by the distribution of borrowers' financial conditions, which takes time to recover.

1.1 Related literature

Our paper is related to the growing body of literature on the relationship between uncertainty (volatility) and economic activities:

One strand of the literature view exogenous increases in uncertainty as a cause of economic recessions.[5] Our paper is related to the recent studies on how financial frictions may amplify the effects uncertainty/volatility shocks on the real economy. For example, Gilchrist et al. (2014) show that financial frictions amplify the effects of uncertainty shocks on investment through variation in credit spreads. Christiano et al. (2014) study a dynamic general equilibrium model incorporated with an agency problem as that in Bernanke et al. (1999), and find that volatility shocks are an important driving force for business cycles. Arellano et al. (2019) consider a model with incomplete financial markets, showing that uncertainty shocks lead to higher default risk and a further decline in hiring. Alfaro et al. (2024) consider a model with both real and financial frictions, finding that the adverse effects of uncertainty shocks are amplified, prolonged and propagated by financial frictions. Ottonello and Winberry (2020) consider the role of financial frictions and firm heterogeneity in the transmission of monetary policy, and find that firms with lower default risk are more

[5] For example, Bernanke (1983), Bloom (2009), Bloom et al. (2018) and Campello et al. (2024) study the real options effects of increased uncertainty, while Gilchrist et al. (2014), Arellano et al. (2019) and Alfaro et al. (2024) focus on how financial frictions may amplify the effects of uncertainty shocks. Gambetti et al. (2025) find that increased uncertainty exerts depressing effects on economic activities only during periods of high agreement.

responsive to monetary shocks. In these studies, stock market volatility is typically viewed as proxies for uncertainty.

In contrast with these studies, our paper considers the reverse causality. Our paper offers a complementary perspective by demonstrating that first-moment productivity shocks can generate counter-cyclical fluctuations in stock market volatility through optimal equity financing contracts under costly state verification and ex-ante monitoring, thereby providing a novel channel by which fundamental shocks translate into financial volatility.

In particular, our model predicts a positive correlation between cross-sectional dispersion of individual stock returns and volatility of market (average) return. This significant empirical pattern is well-documented but may not be clearly explained. Increases in idiosyncratic risks, i.e., risks that can be reduced or eliminated through diversification, can indeed increase cross-sectional dispersion, but not necessarily volatility of market average. If there exists a linkage between aggregate and idiosyncratic risks, what is it? This paper explains this phenomenon from a different perspective: Since the amounts of idiosyncratic and aggregate risks distributed to lenders are both determined by borrowers' financial conditions, thus cross-sectional dispersion and market volatility tend to move together.

Another strand of the literature views higher uncertainty (volatility) as outcomes of economic recessions. Some theories presume that adverse shocks reduce economic activities and learning, leading to less information and thus higher uncertainty (e.g., Van Nieuwerburgh and Veldkamp (2006), Ordoñez (2013) and Fajgelbaum et al. (2017)). Some studies postulate that the increased dispersion and volatility can be caused by adverse first-moment shocks. For example, Gorton and Ordoñez (2014) argue that negative shocks trigger acquisitions of private information on asset quality, leading to larger dispersion of asset values. Ai et al. (2020) argue that negative shocks worsen capital misallocation, leading to counter-cyclical market volatility and cross-sectional dispersion in asset returns. Bachmann and Moscarini (2011) and Tian (2015) argue that negative shocks encourage firms' risk-taking, leading to higher volatility. Our paper complements this literature by providing an alternative explanation why adverse first-moment shocks can cause higher stock market volatility.

Our paper is also related to studies on the dynamic patterns of corporate financing over business cycles. In the earlier studies, such as Bernanke and Gertler (1989), Carlstrom and Fuerst (1997) and Bernanke et al. (1999), there is typically no flexibility in firms' financing arrangement because only debt financing is allowed. More recent studies, such as Hackbarth et al. (2006), Covas and Den Haan (2011, 2012), Jermann and Quadrini (2012) and Begenau and Salomao (2019), consider the dynamic patterns of debt and equity financing over the business cycle.[6] In contrast with those studies, the present paper focuses on external equity financing, and investigates the endogenous variation of the degree of information asymmetry between entrepreneur-managers and outside investors over the business cycle and the implications for stock return volatility.

The rest of the paper is organized as follows. Section 2 describes and solves the model. Section 3 applies the model to account for the observed countercyclical fluctuations in stock return volatility. Section 4 concludes.

2 The Model

We propose a theory that explains the cyclical fluctuations of stock market volatility through the lens of a costly-state-verification (CSV henceforth) framework incorporated with ex-ante monitoring. In the model, the borrower and lenders jointly invest in a project, and realized investment revenues are split between them. The financial contract signed between the borrower and lenders are interpreted as external equity. As will be shown below, volatility of stock return depends on how realized revenues of the project are to be split between the borrower and lenders, which in turn depends on the agency friction. We investigate how the borrowers' financial conditions and thus the severity of agency friction vary over business cycles and the implications for dispersion/volatility of stock returns.

2.1 Environment

Time is discrete and the horizon is infinite, i.e., $t = 0, 1, ...$ There is only one type of goods, that is, consumption goods, which are the 'numerair', and can be consumed or invested. All the prices and quantities are expressed in terms of consumption goods.

[6]Earlier studies, such as Santos (1997), Boyd and Smith (1998, 1999), Chang (1999) and Bolton and Freixas (2000), typically focus on the design of optimal financial contracts and optimal capital structure in the presence of different financial frictions, rather than their variation over the business cycle. For detailed surveys, please see Harris and Raviv (1991) and Franklin and Winton (1995).

There are two groups of agents: borrowers and lenders, with the mass of each group being normalized to one.

Lenders are risk neutral, who can either lend to borrowers or invest in a storage technology that has a constant gross return of R that is exogenously given.

Borrowers, indexed by $i \in [0, 1]$, discount future consumption at rate β, and are also risk neutral. The life-time utility of the borrower is given as $\mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t c_{it}$, where c_{it} denotes consumption in period t.

In period t, borrower i starts with w_{it} units of consumption goods. The wealth distribution is denoted by cumulative density function $\Phi_t(w)$. At the beginning of the period, each borrower is endowed with an investment opportunity, called 'project'. To start the project, the borrower has to invest *one* unit of goods. The part of funds invested by the borrower him/herself is called 'net worth', denoted by n_{it}. The rest of the funds invested, $1 - n_{it}$, is borrowed from lenders. However, as will be discussed below, external financing is subject to information asymmetry.

In what follows, we temporarily drop the borrowers' index i for convenience.

2.1.1 Ex-ante and ex-post monitoring

In our theoretical analysis, an important deviation from conventional CSV models is that we allow for *ex-ante monitoring*, which refers to setting up a mechanism in advance which can credibly force borrowers to reveal the actual investment revenues to lenders. Ex-ante monitoring eliminates the borrower's capability to mis-report profits, and thus reduces the degree of information asymmetry and the costs of external financing.[7]

With ex-ante monitoring, the borrower's investment revenue, y, consists of an *observable* part, denoted by y_z, which can be observed by both the borrower and lender, and an *unobservable* part, denoted by y_u, which is only privately observed by the borrower:

$$y = y_z + y_u, \tag{1}$$

[7] A classical example of ex-ante monitoring is the uses of outside professional auditors (who are familiar with the borrower's project and can interpret any data about the project's return). It became common in the late 19th century in the U.K. (Watts and Zimmerman (1983)).

where

$$\underbrace{y_z = \alpha \cdot e^z \cdot (\rho R_K)}_{observable} \qquad and \qquad \underbrace{y_u = (1-\alpha) \cdot e^u \cdot R_K}_{unobservable}. \tag{2}$$

Here, $\alpha \in [0,1]$ denotes the fraction of investment under ex-ante monitoring, and thus $1-\alpha$ is the part of investment not under ex-ante monitoring. The aggregate investment return, R_K, is observed by the public. z and u are two idiosyncratic shocks. As can be seen from expressions (2), the distributions of y_z and y_u depend on the choice of α. Let $f_{y_z}(y_z)$ and $F_{y_z}(y_z)$ denote respectively the probability density function and cumulative density functions of y_z. Similarly, the p.d.f. and c.d.f. of y_u are denoted by $f_{y_u}(y_u)$ and $F_{y_u}(y_u)$. Importantly, the intensity of ex-ante monitoring, α, is endogenously determined in the optimal financial contract, which may vary with the borrower's dependence on external financing.[8]

Furthermore, ex-ante monitoring is also costly. As can be seen in (2), the expected return of investment under ex-ante monitoring is a fraction $\rho \in (0,1)$ of that of investment not under ex-ante monitoring. The cost of ex-ante monitoring, $1-\rho$, serves as a flexible approach to capture not only the direct costs (e.g., the costs to hire a professional auditor), but also indirect costs associated improving observability of realized profits. For example, during recessions, firms may reduce R&D investment, which may have higher returns but are subject to more severe information asymmetry, when compared to fixed capital investment.[9]

In contrast to ex-ante monitoring, ex-post monitoring takes place only when the project is liquidated by lenders, as in conventional CSV models. For the fraction of investment not under ex-ante monitoring, the realized revenue can only be observed when liquidation (ex-post monitoring) happens. In such cases, lenders have to pay a monitoring cost that is equal to a fraction μ of the unobservable investment revenue.

Obviously, when ex-ante monitoring is not used at all ($\alpha = 0$), the optimal contract is standard debt as in conventional CSV models. When the borrower's investment is under fully ex-ante monitoring ($\alpha = 1$), informa-

[8] As Leland and Pyle (1977) and Myers and Majluf (1984) point out, issuing new equity can be costly due to asymmetric information. New equity issues are usually following additional credible information releases (Korajczyk et al. (1991)). Also, empirical studies have provided plenty of evidence that firms tend to increase information disclosure before and after security issuance (Healy and Palepu (2001)).

[9] See, for example, Brown et al. (2009) and Aghion et al. (2012).

tion asymmetry is entirely eliminated and the optimal contract is standard equity, with payment to lenders being perfectly correlated with realized revenues.

Assumption 1 *The cost of ex-ante monitoring is between zero and the cost of ex-post monitoring, i.e.,* $0 < 1 - \rho < \mu$.

Assumption 1 ensures a trade-off between the two monitoring technologies. When ex-ante monitoring is more costly than ex-post monitoring (i.e., $1-\rho > \mu$), it is never optimal to use ex-ante monitoring. This is because the cost of ex-ante monitoring is incurred with certainty, while the cost of ex-post monitoring is incurred only in the event of liquidation. In this case, the borrower would optimally choose $\alpha = 0$. In contrast, if ex-ante monitoring is free (i.e. $\rho = 1$), the borrower always sets up full ex-ante monitoring (i.e., $\alpha = 1$), which completely eliminates information asymmetry.

Assumption 2 $\rho R_K > R$ *always holds.*

Remember that ρR_K is the expected return of investment under ex-ante monitoring. Thus, Assumption 2 ensures that the surplus from investment under ex-ante monitoring is strictly positive. As will be shown below, this assumption ensures that the borrower's value is always strictly positive.

2.1.2 Shocks

We assume that lenders can only imperfectly infer unobservable revenue y_u by observing the observable part, y_z. Specifically, we assume both shocks z and u comprise a common component and an idiosyncratic component:

$$z = s + \varepsilon_z \qquad \textit{and} \qquad u = s + \varepsilon_u, \tag{3}$$

where s denotes the common component, which captures the impacts of the macro- and industry-level factors, and ε_z and ε_u denote the idiosyncratic components, which are independent from one another, and are i.i.d. across borrowers and over time. Also, ε_z and ε_u are independent from the common component, s. Since for a given borrower, ε_z and ε_u are independent of each other, thus z and u are not perfectly correlated. In fact, z and u are independent conditional on s.

We assume that

$$s \sim \mathcal{N}(-\sigma_s^2/2, \sigma_s^2) \quad \textit{and} \quad \varepsilon_z, \varepsilon_u \sim \mathcal{N}(-\sigma_\varepsilon^2/2, \sigma_\varepsilon^2).$$

Notice that the joint normal distribution of s and z is given by:

$$\begin{bmatrix} s \\ z \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} -\sigma_s^2/2 \\ -(\sigma_s^2 + \sigma_\varepsilon^2)/2 \end{bmatrix}, \begin{bmatrix} \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 + \sigma_\varepsilon^2 \end{bmatrix}\right)$$

Conditional on z, the conditional distribution of s is given by:

$$s|z \sim \mathcal{N}(\kappa z, \sigma_{s|z}^2),$$

where $\kappa \equiv \sigma_s^2/(\sigma_s^2 + \sigma_\varepsilon^2)$ and $\sigma_{s|z}^2 \equiv \sigma_s^2\sigma_\varepsilon^2/(\sigma_s^2 + \sigma_\varepsilon^2)$. Thus, the realized value of z does provide some information on s, but the idiosyncratic component obscure the signal. It follows that the distribution of u conditional on z is

$$u|z \sim \mathcal{N}(\kappa z + \varsigma, \sigma_{u|z}^2). \tag{4}$$

where $\sigma_{u|z}^2 \equiv \sigma_{s|z}^2 + \sigma_\varepsilon^2$ and $\varsigma \equiv -\sigma_\varepsilon^2/2$.

2.2 Financial contract

The lender proposes a menu of contracts, where the contractual terms are contingent on the borrower's net worth n.[10] Then, the borrower determines the level of net worth to invest in the project and selects the corresponding contract.

Given net worth n, the borrower borrows $1 - n$ from the lender. The financial contract, denoted by $\mathcal{A}(n)$, specifies: (1) the intensity of ex-ante monitoring, α, and (2) repayment rules and ex-post verification region.[11]

We focus on incentive-compatible contracts that induce the borrower to truthfully report the realized total revenue y.[12] Since y_z is observable

[10]Appendix A3.2 shows that given the bargaining powers of the lender and the borrower, the optimal contract is invariant to who proposes the contract.

[11]In the benchmark model, we assume that α is perfectly observable and contractible. In Appendix A3.1, we relax this assumption and show that the model results remain qualitatively the same even if α is non-contractible.

[12]Ex-post monitoring happens when the reported value, $\tilde{y}$, falls within the ex-post verification region $D(y_z)$. Following Townsend (1979), we focus on incentive-compatible contracts, in which the borrower always tells the truth. Thus, the reported value is always equal to the actual value, i.e., $\tilde{y} = y$.

to both parties, specified rules can always be made contingent on y_z. Let $x(y, y_z)$ denote the transfer from the borrower to the lender. The ex-post monitoring region, denoted by $D(y_z)$, is a set of realization of y, which also depends on the realized y_z.

Under an incentive-compatible contract, the specified transfer cannot be conditioned on the realized y_u when there is no ex-post verification, as y_u is observable only to the borrower. Further, repayment in the verification region $D(y_z)$ should be no more than that in the non-verification region. Thus, the contract should satisfy:

$$x(y, y_z) = \begin{cases} \overline{x}(y_z) & \textit{if } y \notin D(y_z), \\ \underline{x}(y, y_z) & \textit{if } y \in D(y_z), \end{cases} \tag{5}$$

with

$$\underline{x}(y, y_z) \leq \overline{x}(y_z). \tag{6}$$

where $\overline{x}(y_z)$ represents the repayment in the absence of ex-post verification, while $\underline{x}(y, y_z)$ represents the repayment under verification.

Also, the transfer from the borrower to the lender must be non-negative and satisfy the feasibility constraint:

$$0 \leq x(y, y_z) \leq y. \tag{7}$$

Finally, the borrower should get an expected revenue no less than the required level, $\hat{W}(n)$, which in turn depends positively on how much net worth n the borrower has invested into the project. Thus, the promised return constraint is given as

$$\int_0^\infty \Bigg\{ \underbrace{\int_{y \notin D(z)} \Big(y - \overline{x}(y_z)\Big)\, dF_{y|y_z}(y|y_z; \alpha)}_{\textit{No ex-post monitoring}} + \underbrace{\int_{y \in D(z)} \Big(y - \underline{x}(y, y_z)\Big)\, dF_{y|y_z}(y|y_z; \alpha)}_{\textit{Ex-post monitoring}} \Bigg\}\, dF_{y_z}(y_z; \alpha) \geq \underbrace{\hat{W}(n)}_{\substack{\textit{Required} \\ \textit{revenue}}}. \tag{8}$$

On the left hand side of the above condition, the first term is the borrower's expected revenue if ex-post monitoring does not happen, and the second term is the expected revenue if ex-post monitoring happens.

Problem 1 (Optimal contract I) *Given the borrower's net worth n, the optimal financial contract is the solution of the following optimization problem:*

$$\hat{\Pi}(n) = \max_{\mathcal{A}(n)} \int_0^\infty \Bigg\{ \underbrace{\int_{y \notin D(z)} \overline{x}(y_z)\, dF_{y|y_z}(y|y_z;\alpha)}_{\text{No ex-post monitoring}}$$

$$+ \underbrace{\int_{y \in D(z)} \Big[\underline{x}(y, y_z) - \mu \cdot (y - y_z)\Big]\, dF_{y|y_z}(y|y_z;\alpha)}_{\text{Ex-post monitoring}} \Bigg\} dF_{y_z}(y_z;\alpha) - \underbrace{(1-n) \cdot R}_{\text{Funding cost}}, \quad (9)$$

subject to (5)-(8).

In expression (9), $\hat{\Pi}(n)$ represents the lender's expected net revenue. On the right hand side, the first term is the expected revenue when ex-post monitoring does not happen. The second term is the revenue when ex-post monitoring happens, where $y - y_z = y_u$ is the unobservable revenue (see equation (1)). The third term is the funding cost.

2.3 Characterization of the optimal contract

We start by characterizing the optimal repayment rule $x(y, y_z)$ and the ex-post verification region $D(y_z)$, conditional on a given value of α (Subsection 2.3.1). Then, we consider the optimal choice of α (Subsection 2.3.2).

2.3.1 Optimal repayment schedule and ex-post monitoring

Note that given the value of α, if the observable revenue y_z is fixed, Problem 1 collapses to a standard CSV problem. In such cases, the optimal contract is a standard debt, with a fixed default threshold y^T and verification region $D = [0, y^T]$. The repayment rule is $\overline{x} = y^T$ in the absence of verification, and $\underline{x} = y$ when verification occurs.

In contrast, in our model the distribution of the total revenue y depends on the realization of y_z. Specifically, conditional on y_z, the total revenue y follows a shifted log-normal distribution, with a support of $[y_z, \infty)$ (see equation (1)). In such cases, a standard debt contract with a constant default threshold y^T is no longer optimal. Instead, it becomes optimal to smooth the costs of ex-post monitoring across different states of y_z by adjusting the threshold based on realized y_z.

We show that given α, the optimal contract resembles a state-contingent debt where the default threshold, $y^T(y_z)$, varies with the observed performance y_z.

Lemma 1 *The optimal contract has the following properties for any given realization of y_z:*

1. $D(y_z) = [0, y^T(y_z)]$, *with* $y^T(y_z) \geq y_z$;
2. $\overline{x}(y_z) = y^T(y_z)$ *and* $\underline{x}(y, y_z) = y$.

Proof. *See Appendix A1.1.* ■

As stated in Lemma 1, conditional on y_z, the optimal contract resembles a 'performance-sensitive' debt contract, where ex-post monitoring occurs when the realized (and reported) total revenue y falls below the threshold, $y^T(y_z)$. Importantly, the threshold, $y^T(y_z)$, is not constant but a function of the observed performance, y_z. The repayment rule is give by $\overline{x}(y_z) = y^T(y_z)$ in the absence of ex-post monitoring and $\underline{x}(y, y_z) = y$ otherwise.

Problem 2 (Optimal contract II) *Applying the results from Lemma 1, Problem 1 simplifies to:*

$$\hat{\Pi}(n) = \max_{\alpha, \{y^T(y_z)\}} \int_0^\infty \left\{ y^T(y_z) \cdot \Big(1 - F_{y|y_z}(y^T(y_z)|y_z;\alpha)\Big) + \int_0^{y^T(y_z)} y \, dF_{y|y_z}(y|y_z;\alpha) \right.$$
$$\left. -\mu \cdot \int_0^{y^T(y_z)} (y - y_z) \, dF_{y|y_z}(y|y_z;\alpha) \right\} dF_{y_z}(y_z;\alpha) - (1-n) \cdot R, \quad (10)$$

subject to

$$\int_0^\infty \left\{ \int_{y^T(y_z)}^\infty \Big(y - y^T(y_z)\Big) \, dF_{y|y_z}(y|y_z;\alpha) \right\} dF_{y_z}(y_z;\alpha) \geq \hat{W}(n). \quad (11)$$

2.3.2 Optimal ex-ante monitoring α

Remember that conditional on the observable revenue y_z, the total revenue y follows a shifted log-normal distribution, with a support of $[y_z, \infty)$ (see equation (1)). It is helpful to define

$$y_u^T(y_z) = y^T(y_z) - y_z, \tag{12}$$

and

$$u^T(z) = \log y_u^T(y_z) - \log \hat{y}_u(\alpha). \tag{13}$$

Here, $\hat{y}_u(\alpha) = (1-\alpha)R_K$ is the expected values of y_u given α (see equation (2)). $y_u^T(y_z)$ is the threshold for unrecoverable revenue y_u and $u^T(z)$ is the threshold for unobservable shock u. When $u > u^T(z)$, the realized total revenue y is higher than the threshold $y^T(y_z)$, and thus ex-post monitoring will not happen. When $u \leq u^T(z)$, ex-post monitoring happens and the lenders liquidate the project, taking away all the residuals.

Substituting (2) and (13) into Problem 2 allows us to express it in terms of and $u^T(z)$ and z, yielding the simplified form:

Problem 3 (Optimal contract III) *Problem 2 can be simplified as follows:*

$$\hat{\Pi}(n) = \max_{\alpha\in[0,1],\{u^T(z)\}} \hat{y}_z(\alpha) + \hat{y}_u(\alpha)\cdot \int_{-\infty}^{\infty} \Big[\Gamma(u^T(z)|z) - \Theta(u^T(z)|z)\Big] dF_z(z) - (1-n)\cdot R,$$

subject to

$$\hat{y}_u(\alpha)\cdot\left(1 - \int_{-\infty}^{\infty} \Gamma(u^T(z)|z) dF_z(z)\right) = \hat{W}(n), \tag{14}$$

where $\hat{y}_z(\alpha) = \alpha\cdot(\rho R_K)$ is the expected value of y_z given α, $\hat{y}_u(\alpha) = (1-\alpha)\cdot R_K$ is the expected values of y_u, and $\Gamma(u^T(z)|z)$ and $\Theta(u^T(z)|z)$ are defined as

The lenders' share: $$\Gamma(u^T(z)|z) = e^{u^T(z)}[1 - F_{u|z}(u^T(z)|z)] + \int_{-\infty}^{u^T(z)} e^u f_{u|z}(u|z) du,$$

Liquidation costs: $$\Theta(u^T(z)|z) = \mu \int_{-\infty}^{u^T(z)} e^u f_{u|z}(u|z) du.$$

Conditional on z, $\Gamma(u^T(z)|z)$ is the expected share of unobservable revenue distributed to lenders, and $\Theta(u^T(z)|z)$ is the expected cost of liquidation (ex-post monitoring). Remember that $u^T(z)$ is the threshold of liquidation conditional on z, and $f_{u|z}(u|z)$ and $F_{u|z}(u|z)$ are respectively the p.d.f. and c.d.f. of u conditional on z (see (4)).

As can be seen, the simplified version of the optimal contracting problem (Problem 3) shares some similarities with those in conventional CSV models (e.g., Bernanke et al. (1999)). Nevertheless, our model differs from conventional models in some important ways. In our model, a portion of investment revenue is observable (due to ex-ante monitoring), and the fraction of observable revenue, α, is optimally chosen by the lender. Further, the threshold for ex-post monitoring, $u^T(z)$, is contingent on the observed performance, z. As will be discussed further below, this state-dependency allows the lender to smooth ex-post monitoring costs across different realizations of z.

Optimal threshold $u^T(z)$. The first order condition for $u^T(z)$ is given by:

$$\lambda = \frac{\Gamma'(u^T(z)|z) - \Theta'(u^T(z)|z)}{\Gamma'(u^T(z)|z)} \quad \textit{for all } z.$$

where λ is the Lagrangian multiplier associated with the promised return constraint (14). Notice that on the right hand side of the above condition, the numerator is the marginal increase in the expected net payment to the lender when raising threshold $u^T(z)$, which equals to the marginal increase in the gross payment, $\Gamma'(u^T(z)|z)$, net of the marginal increase in the cost of ex-post monitoring, $\Theta'(u^T(z)|z)$; The denominator is the marginal increase in the gross payment, $\Gamma'(u^T(z)|z)$. Thus, λ is interpreted as the *marginal rate of transformation*: To increase the expected net payment to the lender by one unit, the borrower's expected revenue will be reduced by $1/\lambda$ units, where $1/\lambda - 1$ is the loss rate of transformation due to monitoring costs. It can be shown that the above condition can be re-written as

$$\lambda = 1 - \mu \cdot \Omega(u^T(z) - \kappa z) \quad \textit{for all } z,$$

where function $\Omega(\cdot)$ is defined as $\Omega(a) = \tilde{f}(a)/(1 - \tilde{F}(a))$, with $\tilde{f}(\cdot)$ and $\tilde{F}(\cdot)$ being respectively the p.d.f. and c.d.f. of $\mathcal{N}(\varsigma, \sigma^2_{u|z})$.[13] Notice that on the left hand side of the equation, λ, does not depend on z, and also that on the right hand side, $\Omega(\cdot)$ is a monotonically increasing function.[14] Thus, $u^T(z) - \kappa z$ must equal across state z. This implies

$$u^T(z) = u^* + \kappa z, \tag{15}$$

[13] See Appendix A1.2 for the derivation of $1 - \mu \cdot \Omega(u^T(z) - \kappa z)$.

[14] See Appendix A1.3 for the properties of function $\Omega(\cdot)$.

and

$$\lambda = 1 - \mu \cdot \Omega(u^*), \tag{16}$$

In expression (15), u^* is referred to as the *baseline threshold for ex-post monitoring*, and κz is the optimal adjustment of the threshold according to the realization of z (the observed performance) to smooth ex-post monitoring (liquidation) costs across state z.

Optimal ex-ante monitoring α. Using (15), the first order condition for α can be written as:

$$\rho = \lambda \cdot (1 - \hat{\Gamma}(u^*)) + (\hat{\Gamma}(u^*) - \hat{\Theta}(u^*)), \tag{17}$$

where

$$\hat{\Gamma}(u^*) \equiv \int_{-\infty}^{\infty} \Gamma(u^* + \kappa z|z) dF_z(z) \quad \textit{and} \quad \hat{\Theta}(u^*) \equiv \int_{-\infty}^{\infty} \Theta(u^* + \kappa z|z) dF_z(z)$$

are respectively the lenders' share in unobservable revenue and the expected cost of ex-post monitoring (liquidation).[15]

Condition (17) indicates that the marginal net return of investment under and not under ex-ante monitoring must be equalized. Note first that the expected return of investment under ex-ante monitoring is ρR_K, corresponding to the left hand side of equation (17), which is publicly observable and thus 'pledgeable'. Note also that the expected return of investment not under ex-ante monitoring is R_K, but only a fraction $\hat{\Gamma}(u^*) - \hat{\Theta}(u^*)$ of which is 'pledgeable' (remember that $\hat{\Gamma}(u^*)$ is the share of unobservable revenue distributed to the lenders and $\hat{\Theta}(u^*)$ is the cost of ex-post monitoring). The remaining fraction, $1 - \hat{\Gamma}(u^*)$, is kept by the borrower, which is 'non-pledgeable' and is thus discounted by the *marginal rate of transform*, λ.

Furthermore, notice that the values of λ and u^* are jointly determined by conditions (16) and (17), which implies that they are functions of parameters:

$$\lambda = \lambda(\rho, \mu, \sigma_s^2, \sigma_\varepsilon^2) \quad \textit{and} \quad u^* = u^*(\rho, \mu, \sigma_s^2, \sigma_\varepsilon^2).$$

Then, the promised return constraint (14) can be re-written as:

$$\alpha = 1 - \frac{\hat{W}(n)}{R_K(1 - \hat{\Gamma}(u^*))}. \tag{18}$$

[15]See Appendix A1.4 for the derivation of $\hat{\Gamma}(u^*)$ and $\hat{\Theta}(u^*)$.

Expression (18) indicates that the optimal intensity of ex-ante monitoring, α, is decreasing in the borrower's required revenue $\hat{W}(n)$. Since $\hat{W}(n)$ depends positively on the borrower's net worth n, it follows that for wealthier borrowers, who require less external financing, the intensity of ex-ante monitoring is lower.

Also, as can seen in expression (18), α will never exceed one, but may hits its lower bound of zero when $\hat{W}(n)$ is too large. If this is the case, there exists a critical value of net worth, $\bar{n}$, such that the non-negative constraint for α binds when $n > \bar{n}$. It is given by:

$$\hat{W}(\bar{n}) = R_K \cdot (1 - \hat{\Gamma}(u^*)). \tag{19}$$

In what follows, we focus on the cases where $n \in (0, \bar{n})$.

Proposition 1 (Optimal contract) *Given the borrower's net worth n, with $n \in (0, \bar{n})$, the optimal contract is characterized as follows. The ex-post verification region is $D(y_z) = [0, y^T(y_z)]$, where the threshold, $y_u^T(y_z)$, is given by*

$$y^T(y_z) = y_z + y_u^T(y_z), \quad \textit{where} \quad y_u^T(y_z) = e^{u^* + \kappa z} \cdot \hat{y}_u(\alpha).$$

Here, u^* is a function of parameters as implied by (16) and (17), and the intensity of ex-ante monitoring, α, is determined by condition (18). The repayment in the absence of verification is $\overline{x}(y_z) = y^T(y_z)$, and the repayment with verification is $\underline{x}(y, y_z) = y$.

The borrower's revenue. Note that under the optimal contract the expected revenue of the lender and the borrower, $\hat{\Pi}(n)$ and $\hat{W}(n)$, are given by

$$\hat{\Pi}(n) = \hat{y}_z(\alpha) + \hat{y}_u(\alpha) \cdot \left(\hat{\Gamma}(u^*) - \hat{\Theta}(u^*)\right) - (1 - n) \cdot R,$$
$$\hat{W}(n) = \hat{y}_u(\alpha) \cdot (1 - \hat{\Gamma}(u^*)).$$

Following Gale and Hellwig (1985), we assume that lenders are perfectly competitive. For any given level of net worth n, lenders compete by offering contracts with higher promised revenue, $\hat{W}(n)$, until their expected profits $\hat{\Pi}(n)$ are driven to zero. The zero-profit condition, $\hat{\Pi}(n) = 0$, indicates that

$$\hat{W}(n) = R_E \cdot \left(n + \frac{\rho R_K - R}{R}\right), \tag{20}$$

where $R_E \equiv R/\lambda$ is the expected return on net worth n.[16] Remember that λ is the *marginal rate of transformation*, which is a function of parameters. Thus, R_E is also a function of parameters, which is larger than the cost of external funds, R.

2.4 Intertemporal optimization

We now consider how the optimal level of net worth, n, is determined. In period t, given initial wealth w_{it}, the borrower invests n_{it} into the project and borrows $1 - n_{it}$ from the lenders by choosing a contract $\mathcal{A}_t(n_{it})$, where the borrower's expected revenue is $\hat{W}(n_{it})$. The reminder of the borrower's wealth is consumed, that is, $c_{it} = w_{it} - n_{it}$. The borrower maximizes the life-time utility, given by

$$\mathbb{E}_t \sum_{h=0}^{\infty} \beta^h c_{i,t+h}.$$

The borrower's intertemporal optimization problem can be formulated recursively, leading to the value function below:

Problem 4 (Intertemporal optimization) *The borrower i solves*

$$V(w|\Phi) = \max_{n \in [0,w]} (w - n) + \beta \cdot \mathbb{E}_t \left[V\Big(W(y, y_z, n) | \Phi' \Big) \right],$$

where

$$W(y, y_z, n) = \begin{cases} y - \overline{x}(y_z), & \text{if } y \notin D(y_z), \\ 0, & \text{otherwise.} \end{cases}$$

Remember that productivity shocks z and u have no persistence, and the only aggregate state is the wealth distribution across borrowers, $\Phi(w)$. The evolution of the borrowers' wealth distribution is denoted by $\Phi(w') = \mathcal{F}(\Phi(w), s')$.

Notice that the borrower's expected revenue, i.e., $\hat{W}(n) = \mathbb{E}_{z,u} W(y, y_z, n)$, given by equation (20), is linear in the borrower's net worth, n, indicating that net worth n has a constant expected return, R_E (recall that $R_E \equiv R/\lambda$ and λ is a function of parameters). To ensure that Problem 4 has interior

[16] See Appendix A1.5 for the derivation of the expression for $\hat{W}(n)$.

solutions with non-zero consumption and internal finance, we made the following assumption.

Assumption 3 *Parameter β satisfies $\beta \cdot R_E = 1$.*

Assumption 4 *Whenever the borrower is indifferent between consumption and investing in his/her project, he/she chooses to invest in the project.*

Under Assumptions 3 and 4, the borrower's optimal financing policy is simple. When the borrower's initial wealth w is less than the threshold $\bar{n}$, the borrower invests all the funds in the project, i.e., $n = w$, and borrows $1 - n$ from lenders. When $w > \bar{n}$, the borrower invests $\bar{n}$ in the project and borrows $1 - \bar{n}$ from lenders, and then consumes $w - \bar{n}$. This is because once n exceeds $\bar{n}$, the optimal ex-ante monitoring intensity, α, hits its zero lower bound, and the marginal return on net worth n drops below $1/\beta$.

Proposition 2 (Optimal internal financing) *Given the borrower's initial wealth w:*

1. *If $w \leq \bar{n}$, the borrower invests all the funds in the project, i.e. $n = w$ and $c = 0$;*

2. *If $w > \bar{n}$, the borrower invests $\bar{n}$ in the project and consumes the rest, i.e. $n' = \bar{n}$ and $c = w - \bar{n}$.*

Thus, a borrower with $w > \bar{n}$ is considered *non-financially constrained* in the sense that his/her consumption in the current period is strictly positive.

Finally, it can be shown that the borrower's value is given as

$$V_t(w_t) = w_t + \mathbb{E}_t \sum_{h=0}^{\infty} \beta^h \left[\rho \left(\frac{R_{K,t+h}}{R} \right) - 1 \right].$$

The first term is the borrower's current wealth, and the second term is the surplus from projects in the future periods, which is strictly positive under Assumption 2.

2.5 Determination of aggregate variables

The aggregate level of effective investment, K, is defined as

$$K = \int_i \rho\alpha_i + (1-\alpha_i).$$

Notice that K depends on borrowers' information decisions, α_i. Remember that with ex-ante monitoring, each unit of consumption goods is transformed into $\rho < 1$ units of effective capital, while without ex-ante monitoring, the transformation rate is one. As mentioned in Section 2.1, the cost of ex-ante monitoring, $1-\rho$, captures not only the direct costs but also the indirect costs associated with improving observability of realized profits.[17]

To close the model, we assume that the aggregate investment return, R_K, decreases with K:

$$R_K = AK^{\theta-1}, \tag{21}$$

where $\theta \in (0,1)$ and A is a constant. One can think that there are goods producers who rent capital from the borrowers and R_K is the rental price. Thus, the aggregate *corporate profits* is given by

$$Y = \int_i \left[\rho\alpha_i e^{z_i} + (1-\alpha_i)e^{u_i}\right] R_K = e^s AK^{\theta}, \tag{22}$$

where as mentioned before e^s is the aggregate shock.[18].

2.6 Internal and external equity

The financial contract in our model is interpreted as external equity, and thus lenders correspond to outside investors of firms' equities, and the borrower corresponds to an entrepreneur-manager who not only operates the firm but also owns internal equity. The borrowers' revenue contains not only the return on (internal) equity but also managerial compensation.

At the beginning of the current period, the amount of internal equity (of the entrepreneur-manager) is n_i, and the amount of external equity (held

[17]For example, to increase observability of investment revenue, a firm may reduce R&D investment, which is more efficient but subject to more severe information asymmetry, when compared to fixed capital investment.

[18]Recall that $z_i = s + \varepsilon_{z,i}$ and $u_i = s + \varepsilon_{u,i}$ where the idiosyncratic components ($\varepsilon_{z,i}$ and $\varepsilon_{u,i}$) are i.i.d. across borrowers and do not depend on borrowers' choices of α_i.

by outside investors) is $1-n_i$, which are determined in the previous period. If the project is not liquidated, the ex-post return of equity in the current period is given as

$$r_i = \left[\frac{\alpha_i \cdot \rho \cdot e^{z_i} + (1-\alpha_i) \cdot e^{u^* + \kappa z_i}}{1-n_i}\right] \cdot R_K. \tag{23}$$

Remember that in the above expression the numerator is the payment to lenders (see Proposition 1), while the denominator is the amount borrowed from lenders. At the end of the current period, the amounts of internal and external equity are respectively n_i' and $1-n_i'$, which will be carried into the next period. Thus, the net external equity financing in the current period (i.e., issuance of additional external equity minus share repurchases and dividends) is given as

$$\textit{Net external equity financing} = \underbrace{(1-n_i') - (1-n_i)}_{\textit{new issues minus repurchases}} - \underbrace{(r_i - 1)(1-n_i)}_{\textit{dividends}}.$$

If the project is liquidated in the current period, the borrower starts a new firm in the next period with zero internal equity.

Before analyzing how the stock market volatility fluctuates over business cycles, it is helpful to formally define these measures.

The likelihood of liquidation. Remember that liquidation happens only when realized u is below the threshold $u^T(z) = u^* + \kappa z$. We denote by $\mathcal{B}(s)$ the set of borrowers whose projects are not liquidated, given the aggregate shock s. Thus, $\mathcal{B}(s)$ is defined as

$$\mathcal{B}(s) = \left\{ i \,| u_i > u^* + \kappa z_i \right\}.$$

The share of projects that are liquidated is

$$1 - \int_{i \in \mathcal{B}(s)} 1.$$

Recall that $z_i = s + \varepsilon_{z,i}$ and $u_i = s + \varepsilon_{u,i}$. Thus, the likelihood of liquidation can be written as

$$Prob\left[\varepsilon_{u,i} - \kappa \varepsilon_{z,i} < u^* - (1-\kappa)s\right].$$

Remember that $\varepsilon_{z,i}$ and $\varepsilon_{u,i}$ are independent of s and $\kappa < 1$. Thus, when the aggregate productivity, s, is low, a larger share of projects will be liquidated.

The cross-sectional dispersion in equity returns. The cross-sectional standard deviation in ex-post stock returns, denoted by σ_p, is defined by

$$\sigma_p^2 = \frac{\int_{i \in \mathcal{B}(s)} (r_i - \bar{r})^2}{\int_{i \in \mathcal{B}(s)} 1}. \tag{24}$$

where $\bar{r}$ is the average of r_i, with $i \in \mathcal{B}(s)$.

As can be seen from expressions (23) and (24), the cross-sectional dispersion of individual equity return r_i depends not only on the intensity of ex-ante monitoring, but also on the heterogeneity in borrowers' financing decisions. Notice that equity return r_i is more responsive to the observed performance, z_i, when the intensity of ex-ante monitoring, α_i, is higher. Also, the heterogeneity in borrowers' choices of α_i and n_i also contribute to the cross-sectional dispersion of equity returns.

The volatility of market return. The market return, r_m, is defined as

$$r_m = \int_{i \in \mathcal{B}(s)} \omega_i r_i, \tag{25}$$

where ω_i is the weight of borrower i, defined as

$$\omega_i = \frac{1 - n_i}{\int_{i \in \mathcal{B}(s)} (1 - n_i)}.$$

We denote by σ_m^2 the ex-ante variance of market return, which is given as

$$\sigma_m^2 = \mathbb{E}(r_m^2) - \mathbb{E}(r_m)^2. \tag{26}$$

The expression for σ_m^2 is derived in Appendix A1.6.

Financial condition. The economy-wide financial condition is captured by the spread between the aggregate investment return, R_K, and the funding cost, R:

$$Spread = R_K - R.$$

3 Cyclical dispersion and volatility

Stock market volatility measures, such as cross-sectional dispersion of firm-level stock returns and volatility of market return, tend to move together in a counter-cyclical manner. Specifically, empirical studies have found:

1. The cross-sectional dispersion and market volatility in stock returns fluctuate counter-cyclically and are highly correlated with each other (e.g., Campbell and Lettau (1999) and Bloom (2009)).

2. Stock market volatility measures are highly correlated with financial condition indicators (e.g., Gilchrist et al. (2014)).

These facts are often used as evidence supporting the notion that uncertainty shocks (i.e., exogenous increases in uncertainty) lead to economic recessions. However, our theoretical analysis in the previous sections suggests that this phenomenon can be a byproduct of optimal financing adjustment following adverse first-moment shocks. Here, we use the calibrated version of the model to quantitatively show that a negative first-moment shock to the aggregate productivity can lead to increases in stock market volatility through the deteriorating the financial conditions of firms.

3.1 Parameter values

We calibrate the model to monthly frequency. The return of storage technology, R, is set to 1.002, indicating an annualized risk-free rate of 0.024. Capital share, θ, is set to 0.25. Scalar A in equation (21) governs the aggregate level of investment return and also the overall equity return. Its value is chosen such that the annualized average return on equity (conditional on non-liquidation) across firms is around 0.06. The cost of liquidation (ex-post monitoring), μ, is set to 0.2, which is inline with existing studies.[19]

The key parameters are ρ, σ_s and σ_ε. Their values are jointly chosen so that the key model moments are inline with their data counterparts. Notice that parameter σ_ε (i.e., the standard deviation of firm-level idiosyncratic shock ε) governs the degree of cross-sectional dispersion of stock returns, $\sigma_{p,t}$; and parameter σ_s (i.e., the standard deviation of aggregate shock s) governs the overall volatility of market return, $\sigma_{m,t}$. We set $\sigma_s = 0.1$ and

[19] For example, Bernanke et al. (1999) and Carlstrom and Fuerst (1997).

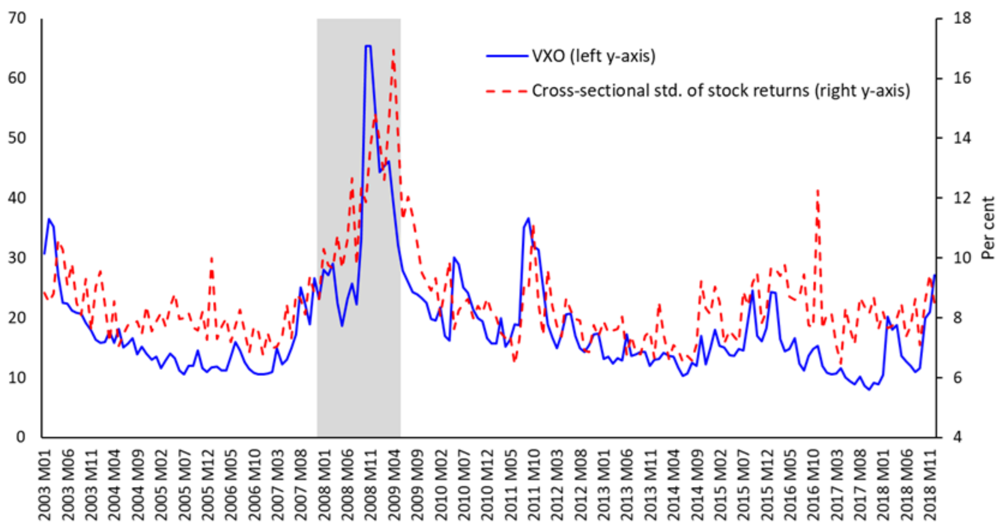

Figure 1: The cross-sectional dispersion and time-series volatility in stock returns

Notes: The above figure shows the movements of the cross-sectional standard deviation of individual stock returns and the market volatility measured by the VXO index over the period 2003M01-2018M12. Shaded bars indicate official NBER recessions.

$\sigma_\varepsilon = 0.1$, and as a result, the long-run means of $\sigma_{p,t}$ and $\sigma_{m,t}$ at the stochastic steady-state can roughly match the actual values before the 2008 financial crisis.[20] Recall that $1-\rho$ is the cost of ex-ante monitoring, which governs the cost of external equity financing. Thus, the value of ρ is chosen to ensure that the average share of external equity in total equity at the stochastic steady state matches the data.[21] This indicates that $1-\rho = 0.024$.[22]

[20] The values chosen for parameters σ_s and σ_ε are also inline with the estimates of Bloom (2009).

[21] According to Kole (1995), around 13% of equity of U.S. listed firms are held by the management.

[22] Also, under our calibration, the cost of issuing external equity, λ, is around 0.1, which is inline with the estimates of Belo et al. (2019).

Table 1: Parameter values

	Value	Description
R	1.002	Return of the storage technology (monthly)
θ	0.25	Capital share
μ	0.2	Ex-post monitoring cost
ρ	0.976	Ex-ante monitoring cost, $1-\rho$, is 0.024
σ_s	0.1	Aggregate risk
σ_ε	0.1	Idiosyncratic risk

3.2 Correlations

Our analysis focuses on three measures, namely the cross-sectional standard deviation of individual equity returns $\sigma_{p,t}$, the volatility of market return $\sigma_{m,t}$, and the economy-wide financial conditions.

Table 2 below shows the correlation coefficients between pairs of these measures. For each pair, the first columns show the correlation coefficients estimated using corresponding data counterparts, and the second columns show the coefficients estimated using model simulated data. As shown in the table, simulated values are in line with data counterparts.

In the model, the correlation coefficient between market volatility and financial condition is close to one, which is due to the set up of the model. Recall that in the model, financial condition is represented by the spread between aggregate investment return and funding cost, i.e. $R_K - R$, where R_K is endogenously determined by how borrowers allocate funds between investments with different degrees of transparency, that is, the average α. The volatility of market return also depends on the average α, as shown in Section 2.6. As a result, the two measures are almost perfectly correlated.

Table A1 in Appendix A2.1 shows that both the two stock market volatility measures, i.e., cross-sectional standard deviation of individual equity returns $\sigma_{p,t}$ and volatility of market return $\sigma_{m,t}$, are negatively correlated with lagged TFP shocks, which are inline with the data. Remember that in our model, an aggregate TFP shock affects the borrowers' net worth and thus optimal financial contracting in the current period, which then affects the volatility of asset returns in the next period.

Robustness. Figure A1 in Appendix A2.2 shows the correlation coefficient between cross-sectional standard deviation of stock returns, $\sigma_{p,t}$, and

Table 2: Correlations

	I. Cross-sectional standard deviation of stock returns				II. Volatility of market return	
	Data (1)	Model (2)	Data (3)	Model (4)	Data (5)	Model (6)
Financial condition	0.795***	0.719***			0.842***	0.999***
	(0.044)	(0.007)			(0.039)	(0.0004)
Volatility of market return			0.719***	0.720***		
			(0.050)	(0.007)		
R^2	0.632	0.517	0.516	0.519	0.709	0.999
Observations	192	-	192	-	192	-

Notes: The basket of selected stocks are consistent with Bloom (2009), and market volatility is measured using the VXO index. The actual tightness of the economy-wide financial condition is measured using the Chicago Fed National Financial Condition Index. The regressions are conducted using actual data over 2003M01-2018M12 and model-simulated data. All variables are normalized to have a standard deviation of one. Standard errors are reported in the brackets below.

volatility of market return, $\sigma_{m,t}$, computed over a rolling window. The correlation coefficient is within the range of $[0.5, 0.8]$ most of the time. Thus, the positive correlation between the two stock market volatility measures seems robust across different periods.

Tables A2 and A3 in Appendix A2.2 confirm that under our calibration, the key moments of the model are close to their data counterparts. Further, we also compute these moments (at the stochastic steady state) for different values of ρ, σ_s, and σ_ε. As shown in the tables, the results are qualitatively the same as (and quantitatively similar to) those of the benchmark model. Thus, our results are robust to alternative values of these parameters.[23]

[23] As show in the tables, when uncertainty increases (higher σ_s and σ_ε), the correlation coefficient between $\sigma_{p,t}$ and $\sigma_{m,t}$ becomes slightly smaller. This is due to the construction of the two measures: In the model, $\sigma_{p,t}$ (i.e, the standard deviation of realized returns) is an ex-post measure, while $\sigma_{m,t}$ (i.e., the square root of variance of market return) is an ex-ante measure, computed before the realization of current period shocks. Therefore, the correlation between the two measures becomes weaker when uncertainty is higher.

3.3 Numerical exercises

3.3.1 Impulse responses

Recall that productivity shocks z and u are i.i.d. over time with no persistence, and the only aggregate state is the wealth distribution across borrowers, $\Phi(w)$. We assume that the economy is initially at the non-stochastic steady state, and simulate the outcomes using independent Gaussian draws for the model's innovations for twenty thousands replications. Then, we compute resultant impulse responses using the method proposed by Koop et al. (1996). The details are provided in Appendix A1.7.

We consider the effects of a negative first-moment shock that exogenously reduces the aggregate productivity, s, by $2\sigma_s$. Figure 2 below shows the responses of the economy to the shock.

As shown in the Figure 2, following the adverse shock, the borrower's internal equity (net worth) drops. As a result, the agency problem becomes more severe because the borrower now has a smaller stake in the project, leading to higher costs of external financing. To mitigate the increased external financing costs, ex-ante monitoring is used more intensively, which distributes more aggregate and idiosyncratic risks to the lenders, leading to rises in both cross-sectional dispersion of individual stock returns and volatility of market return. Further, the increased intensity of ex-ante monitoring reduces aggregate effective investment, K.[24] This widens the spread between aggregate investment return R_K and funding cost R, reflecting increased economy-wide external financing costs. In sum, our model predicts (i) counter-cyclical movements of cross-sectional dispersion of stock returns and volatility of market return and co-movement between the two measures, and (ii) co-movements between stock market volatility measures and the economy-wide external financing costs, which are consistent with data (Table 2).

In addition, heterogeneity in the borrowers' choices of monitoring technologies, $\{\alpha_i\}$, also increases after the shock. This is because non-financially constrained borrowers always make the same financial and information decisions regardless of their wealth levels, while financially constrained borrowers make different decisions according to their wealth levels. Fol-

[24] As mentioned in Section 2.1, in our model the cost of ex-ante monitoring captures not only the direct costs (e.g., the costs to hire a professional auditor), but also indirect costs associated with improving observability of realized profits.

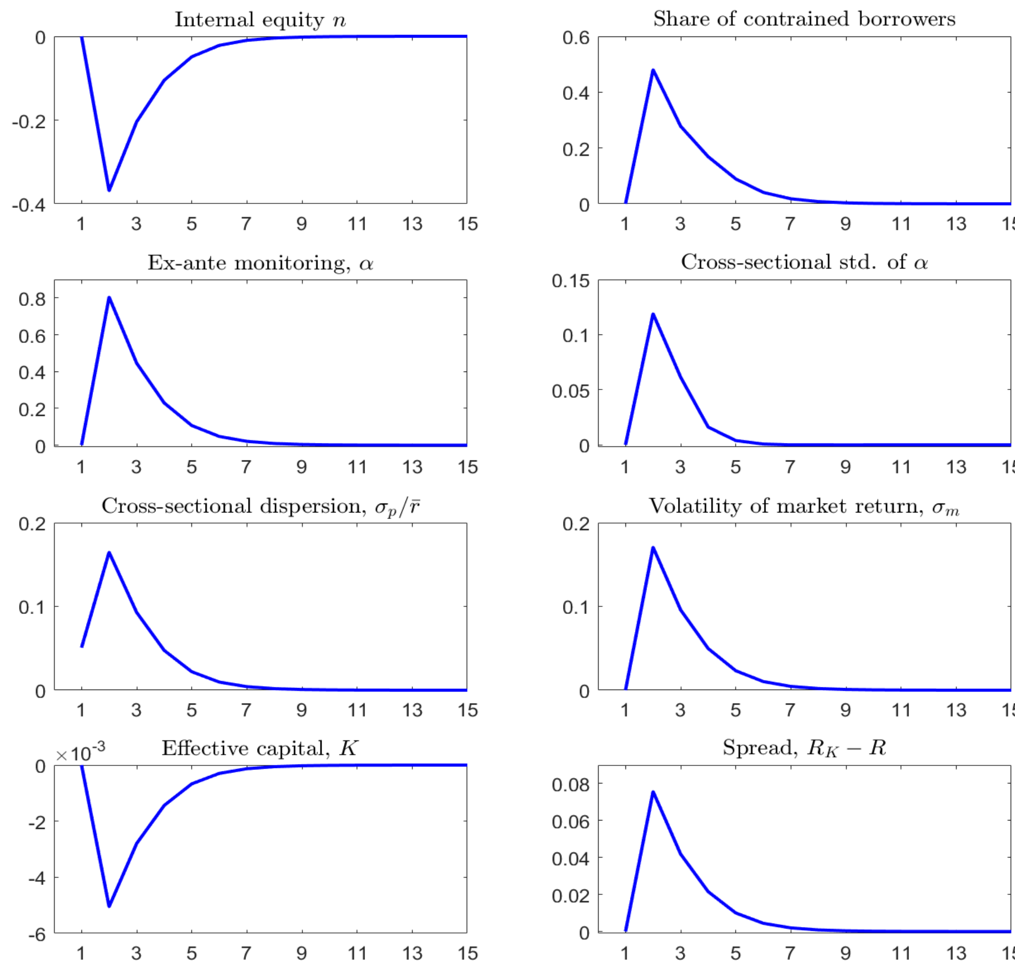

Figure 2: Responses to a negative aggregate TFP shock

Notes: All variables are normalized by corresponding long-run unconditional means. Time horizon in months.

lowing the negative shock, the borrowers' wealth distribution is shifted to the left and more borrowers become financially constrained. Thus, the increased heterogeneity across borrowers also contributes to the increased cross-sectional dispersion of stock returns.[25]

[25]As shown in Figure A4 in Appendix A1.7, the increased heterogeneity across borrowers also leads to larger cross-sectional dispersion of expected output.

Finally, note that all the responses are persistent even if the shocks do not have any persistence at all. The reason is that the stock market volatility measures and financial condition measures considered in our exercises are largely determined by the distribution of borrowers' wealth, and it takes time for the wealth distribution to recover after the negative first-moment shock.

Policy implications. We consider first how financial transparency may affect model results. We define financial transparency as the intensity and quality of financial disclosures (Bushman et al. (2004)). Thus, improvement of financial transparency, such as changes in mandated disclosure, can reduce the information asymmetry between mangers and outside investors, facilitating external monitoring of managers (Berger and Hann (2003)). Thus, in our model a higher level of transparency implies lower costs of both ex-ante and ex-post monitoring. Thus, to investigate the effects of transparency, we set monitoring costs, μ and $1-\rho$, 25% lower than their baseline values and then compute the impulse responses to a negative first-moment shock. The results are shown in Figure A2 in Appendix A2.3.

As shown in Figure A2, with a higher level of financial transparency (i.e., lower ex-ante and ex-post monitoring costs), external financing costs (the spread, $R_K - R$) increase less after the adverse shock; However, stock market volatility increases as much as in the benchmark model. Thus, in our model transparency helps mitigate the fluctuations of external financing costs over business cycles, but does not significantly reduce the cyclical fluctuations of stock market volatility.

This is because in the model, stock market volatility is largely determined by the choice of monitoring technologies: Volatility would be higher if ex-ante monitoring is used more intensively. Reducing the costs of both ex-ante and ex-post monitoring does not change the 'relative cost' of the two, and thus has little impact the optimal choice of monitoring technologies. Thus, the response of stock market volatility is similar to those in the benchmark model. However, with lower monitoring costs, the fluctuations of external financing costs are reduced.

Also, we show that an expansionary monetary stimulus that lowers the real interest rate can effectively mitigate the increases in external financing costs and stock market volatility following the adverse first-moment shock. The impulse responses are shown in Figure A3 in Appendix A2.3. This is because when the real interest rate is lower, the expected payment required by outside investors is less, which reduces the need for monitoring. This result is consistent with empirical studies, such as Bekaert et al. (2013), who found a strong positive correlation between real federal funds rate and the VIX index.

3.3.2 The 2008 financial crisis

To investigate the model's capability to explain the increased volatility during crises, we simulate the model using actual shocks over 2003M01-2018M12. We ask the following question: Can exogenous decreases in aggregate TFP that generate the observed drops in corporate profits during the 2008 financial crisis lead to sizable rises in stock market volatility?

To this end, we first obtain the quarterly series of actual TFP by using the series of quarterly TFP growth constructed by Fernald (2014).[26] Then, we convert the quarterly series to monthly through liner interpolation. TFP shocks are the deviations of TFP from time trend, which are normalized to have a standard deviation of σ_s.

We simulate the model using the actual TFP shocks, while assuming the economy is initially at the non-stochastic steady state. The results are shown in Figure 3, in which all the values are shown in log deviation from their levels at the beginning of the crisis (2007M08). As the figure shows, the model captures the drop in corporate profits. This is unsurprising since actual TFPs are constructed using information on income and expenditure of the business sector. However, the model is also able to explain around 1/2 to 2/3 of the fluctuations in stock market volatility measures during the crisis. Thus, the channel studied in the present paper can be an alternative explanation for the increased stock market volatility during the crisis.

4 Conclusion

The countercyclical fluctuations of stock market volatility is often used as evidence supporting the notion that uncertainty shocks (i.e., exogenous increases in uncertainty) lead to economic recessions. However, we show

[26]The quarterly series of actual TFP growth are constructed using the income and expenditure sides of the national income and products accounts (NIPA), which are available only at quarterly frequency.

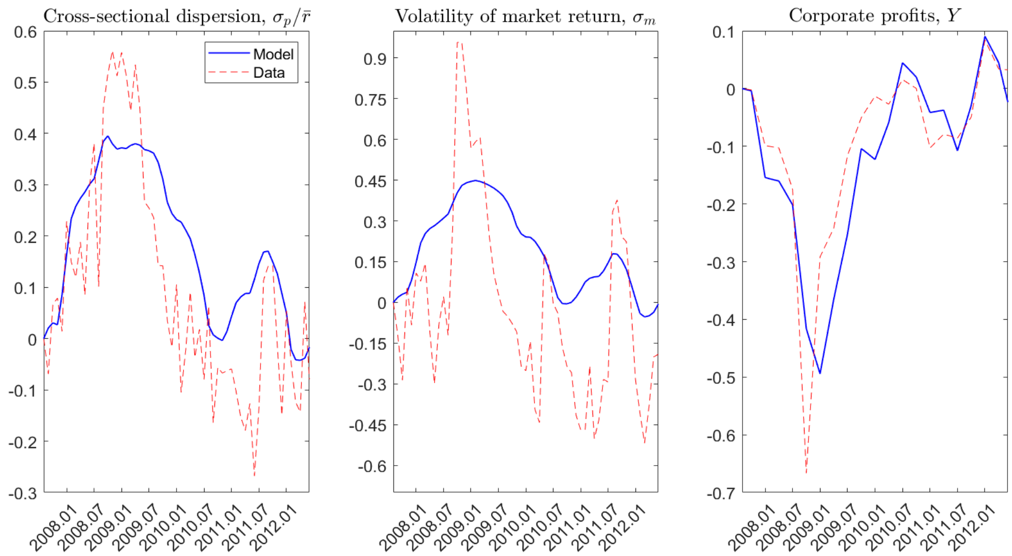

Figure 3: Stock market volatility and corporate profits during the 2008 financial crisis

Notes: The solid-blue lines show the cross-sectional relative dispersion of stock returns, $\sigma_{p,t}/\bar{r}_t$, the volatility of market return, $\sigma_{m,t}$, and aggregate corporate profits, Y_t, during the 2008 financial crisis, which are simulated by using the model and the series of actual TFP shocks constructed by Fernald (2014) (converted from quarterly to monthly through linear interpolation). The dashed-red lines show their data counterparts. All variables are shown in log deviation from their values at the beginning of the crisis (2007M08).

that first-moment productivity shocks can lead to countercyclical fluctuations of measures of stock market volatility through their impacts on firms' financial conditions.

We incorporate a CSV problem into a dynamic framework. The model deviates from conventional CSV models by allowing for ex-ante monitoring, which refers to setting up a mechanism in advance that can credibly force borrowers to disclose the true investment outcomes to lenders. Ex-ante monitoring is costly, but it allows payment to be made contingent on the information revealed, which helps mitigate information frictions and lower the likelihood of costly liquidation (ex-post monitoring).

The financial contracts in our model are interpreted as external equities, and borrowers correspond to entrepreneur-managers who not only operate projects but also own internal equities. As an adverse first-moment shock reduces the value of internal equities, the costs of external equity financing become higher. Consequently, ex-ante monitoring is used more intensively to mitigate information frictions. This facilitates external equity financing, but it also leads to endogenous increases in equity return volatility, as a byproduct of optimal financing adjustment.

Our model predicts countercyclical movements of stock market volatilities and their close correlation with borrowers/firms' financial conditions. The calibrated model can generate correlation coefficients between pairs of these measures that align quantitatively with empirical data.

The results of our analysis indicate that the observed countercyclical fluctuations of stock market volatility can be outcomes of first-moment shocks to fundamentals. However, our results do not necessarily diminish the importance of stock market volatility measures as uncertainty proxies. The primary goal in this paper is to explore another possible explanation for this phenomenon.

References

P. Aghion, P. Askenazy, N. Berman, G. Cette, and L. Eymard. Credit constraints and the cyclicality of R&D investment: Evidence from France. *Journal of the European Economic Association*, 10(5):1001–1024, 2012.

H. Ai, K. Li, and F. Yang. Financial intermediation and capital reallocation. *Journal of Financial Economics*, 138(3):663–686, 2020.

I. Alfaro, N. Bloom, and X. Lin. The finance uncertainty multiplier. *Journal of Political Economy*, 132(2):577–615, 2024.

C. Arellano, Y. Bai, and P. J. Kehoe. Financial frictions and fluctuations in volatility. *Journal of Political Economy*, 127(5):2049–2103, 2019.

R. Bachmann and G. Moscarini. Business cycles and endogenous uncertainty. Working Paper, 2011.

S. Basu and B. Bundick. Uncertainty shocks in a model of effective demand. *Econometrica*, 85(3):937–958, 2017.

J. Begenau and J. Salomao. Firm financing over the business cycle. *Review of Financial Studies*, 32(4):1235–1274, 2019.

G. Bekaert, M. Hoerova, and M. L. Duca. Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788, 2013.

F. Belo, X. Lin, and F. Yang. External equity financing shocks, financial flows, and asset prices. *Review of Financial Studies*, 32(9):3500–3543, 2019.

P. G. Berger and R. Hann. The impact of sfas no. 131 on information and monitoring. *Journal of Accounting Research*, 41(2):163–223, 2003.

B. S. Bernanke. Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics*, 98(1):85–106, 1983.

B. S. Bernanke and M. Gertler. Agency costs, net worth, and business fluctuations. *American Economic Review*, pages 14–31, 1989.

B. S. Bernanke, M. Gertler, and S. Gilchrist. The financial accelerator in a quantitative business cycle framework. *Handbook of Macroeconomics*, 1: 1341–1393, 1999.

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

N. Bloom, M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry. Really uncertain business cycles. *Econometrica*, 86(3):1031–1065, 2018.

P. Bolton and X. Freixas. Equity, bonds, and bank debt: Capital structure and financial market equilibrium under asymmetric information. *Journal of Political Economy*, 108(2):324–351, 2000.

J. H. Boyd and B. D. Smith. The evolution of debt and equity markets in economic development. *Economic Theory*, 12:519–560, 1998.

J. H. Boyd and B. D. Smith. The use of debt and equity in optimal financial contracts. *Journal of Financial Intermediation*, 8(4):270–316, 1999.

J. R. Brown, S. M. Fazzari, and B. C. Petersen. Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *Journal of Finance*, 64(1):151–185, 2009.

R. M. Bushman, J. D. Piotroski, and A. J. Smith. What determines corporate transparency? *Journal of Accounting Research*, 42(2):207–252, 2004.

J. Y. Campbell and M. Lettau. Dispersion and volatility in stock returns: An empirical investigation. Working Paper, National Bureau of Economic Research, 1999.

J. Y. Campbell, M. Lettau, B. Malkiel, and Y. Xu. Idiosyncratic equity risk two decades later. *Critical Finance Review*, 12(1-4):203–223, 2023.

M. Campello, G. Kankanhalli, and H. Kim. Delayed creative destruction: How uncertainty shapes corporate assets. *Journal of Financial Economics*, 153:103786, 2024.

C. T. Carlstrom and T. S. Fuerst. Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review*, pages 893–910, 1997.

A. Cesa-Bianchi, M. H. Pesaran, and A. Rebucci. Uncertainty and economic activity: A multicountry perspective. *Review of Financial Studies*, 33(8): 3393–3445, 2020.

C. Chang. Capital structure as optimal contracts. *North American Journal of Economics and Finance*, 10(2):363–385, 1999.

L. J. Christiano, R. Motto, and M. Rostagno. Risk shocks. *American Economic Review*, 104(1):27–65, 2014.

F. Covas and W. J. Den Haan. The cyclical behavior of debt and equity finance. *American Economic Review*, 101(2):877–899, 2011.

F. Covas and W. J. Den Haan. The role of debt and equity finance over the business cycle. *Economic Journal*, 122(565):1262–1286, 2012.

P. D. Fajgelbaum, E. Schaal, and M. Taschereau-Dumouchel. Uncertainty traps. *Quarterly Journal of Economics*, 132(4):1641–1692, 2017.

J. Fernald. A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco, 2014.

A. Franklin and A. Winton. Corporate financial structure, incentives and optimal contracting. *Handbooks in Operations Research and Management Science*, 9:693–720, 1995.

D. Gale and M. Hellwig. Incentive-compatible debt contracts: The one-period problem. *The Review of Economic Studies*, 52(4):647–663, 1985.

L. Gambetti, D. Korobilis, J. D. Tsoukalas, and F. Zanetti. Agreed and disagreed uncertainty. Working Paper, 2025.

S. Gilchrist, J. W. Sim, and E. Zakrajšek. Uncertainty, financial frictions, and investment dynamics. Working Paper, National Bureau of Economic Research, 2014.

G. Gorton and G. Ordoñez. Collateral crises. *American Economic Review*, 104(2):343–378, 2014.

D. Hackbarth, J. Miao, and E. Morellec. Capital structure, credit risk, and macroeconomic conditions. *Journal of Financial Economics*, 82(3):519–550, 2006.

M. Harris and A. Raviv. The theory of capital structure. *Journal of Finance*, 46(1):297–355, 1991.

P. M. Healy and K. G. Palepu. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1):405–440, 2001.

U. Jermann and V. Quadrini. Macroeconomic effects of financial shocks. *American Economic Review*, 102(1):238–271, 2012.

K. Jurado, S. C. Ludvigson, and S. Ng. Measuring uncertainty. *American Economic Review*, 105(3):1177–1216, 2015.

S. R. Kole. Measuring managerial equity ownership: A comparison of sources of ownership data. *Journal of Corporate Finance*, 1(3-4):413–435, 1995.

G. Koop, M. H. Pesaran, and S. M. Potter. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147, 1996.

R. A. Korajczyk, D. J. Lucas, and R. L. McDonald. The effect of information releases on the pricing and timing of equity issues. *Review of Financial Studies*, 4(4):685–708, 1991.

S. Krasa and A. P. Villamil. Optimal contracts when enforcement is a decision variable. *Econometrica*, 68(1):119–134, 2000.

H. E. Leland and D. H. Pyle. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance*, 32(2):371–387, 1977.

S. C. Ludvigson, S. Ma, and S. Ng. Uncertainty and business cycles: Exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*, 13(4):369–410, 2021.

S. C. Myers and N. S. Majluf. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13(2):187–221, 1984.

G. Ordoñez. The asymmetric effects of financial frictions. *Journal of Political Economy*, 121(5):844–895, 2013.

P. Ottonello and T. Winberry. Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, 88(6):2473–2502, 2020.

J. C. Santos. Debt and equity as optimal contracts. *Journal of Corporate Finance*, 3(4):355–366, 1997.

J. K. Seward. Corporate financial policy and the theory of financial intermediation. *Journal of Finance*, 45(2):351–377, 1990.

C. Tian. Riskiness, endogenous productivity dispersion and business cycles. *Journal of Economic Dynamics and Control*, 57:227–249, 2015.

R. M. Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265–293, 1979.

S. Van Nieuwerburgh and L. Veldkamp. Learning asymmetries in real business cycles. *Journal of Monetary Economics*, 53(4):753–772, 2006.

R. L. Watts and J. L. Zimmerman. Agency problems, auditing, and the theory of the firm: Some evidence. *Journal of Law and Economics*, 26(3): 613–633, 1983.

Appendices

A1 Mathematical derivations and proofs

A1.1 Proof of Lemma 1

Suppose the promised return constraint (8) binds. Then, Problem 1 can be re-written as:

$$\hat{\Pi}(n) = \max_{\mathcal{A}(n)} \quad \hat{y}(\alpha) - \hat{W}(n)$$
$$- \mu \cdot \int_0^\infty \Bigg[\underbrace{\int_{y \in D(z)} (y - y_z)\, dF_{y|y_z}(y|y_z;\alpha)}_{\textit{Ex-post monitoring costs}} \Bigg] dF_{y_z}(y_z;\alpha) - (1-n)\cdot R, \quad \text{(A.1)}$$

subject to

$$\hat{y}(\alpha) - \int_0^\infty \Bigg\{ \underbrace{\int_{y \notin D(z)} \overline{x}(y_z)\, dF_{y|y_z}(y|y_z;\alpha)}_{\textit{No ex-post monitoring}}$$
$$+ \underbrace{\int_{y \in D(z)} \underline{x}(y, y_z)\, dF_{y|y_z}(y|y_z;\alpha)}_{\textit{Ex-post monitoring}} \Bigg\} dF_{y_z}(y_z;\alpha) = \underbrace{\hat{W}(n)}_{\substack{\textit{Required}\\ \textit{revenue}}}, \quad \text{(A.2)}$$

$$\underline{x}(y, y_z) \le \overline{x}(y_z), \quad \text{(A.3)}$$

and

$$0 \le \underline{x}(y, y_z) \le y \quad \textit{and} \quad 0 \le \overline{x}(y_z) \le y, \quad \text{(A.4)}$$

where $\hat{y}(\alpha) = [\alpha\rho + (1-\alpha)]R_K$ is the expected total revenue and $y - y_z$ is the unobservable revenue (see equations (1) and (2)).[A1]

As can be seen from the above optimization problem, given the choice of α, the optimal repayment rule, $x(y, y_z)$ (i.e., $\overline{x}(y_z)$ and $\underline{x}(y, y_z)$), and verification region $D(y_z)$, should minimize the cost of ex-post monitoring (i.e.,

[A1] Given α, the expected values of y_z and y_u are given respectively by $\hat{y}_z(\alpha) = \alpha\rho R_K$ and $\hat{y}_u(\alpha) = (1-\alpha)R_K$ (see equation (2)).

the third term in (A.1)), subject to the promised return constraint (A.2)), the incentive compatibility constraint (A.3), and the feasibility constraints (A.4).

Recall that with ex-ante monitoring, the total revenue, y, consists of an observable part, y_z, and an unobservable part, y_u (see equation (1)). Conditional on y_z, the total revenue y follows a shifted log-normal distribution, with a support of $[y_z, \infty)$. If for a given y_z, the repayment obligation $\overline{x}(y_z)$ is strictly less than y_z, then the contractual transfer can be fully satisfied by using the observable revenue y_z, and thus ex-post monitoring is no longer needed, i.e., $D(y_z) = \emptyset$, regardless of the realized total revenue y. In contrast, if $\overline{x}(y_z) > y_z$, the entirety of the observable revenue y_z is distributed to the lender. The residual obligation, $\overline{x}(y_z) - y_z$, must then be paid from unobservable revenue y_u. In such cases, ex-post monitoring is still needed.

If the intensity of ex-ante monitoring, α, is high enough such that the expected value of unobservable revenue is less than the promised revenue of the borrower, i.e., $\mathbb{E}[y_u] = (1-\alpha)R_K < \hat{W}(n)$, then ex-post monitoring is no longer needed. This is because to satisfy the promised return constraint (A.2), the lender can simply distribute all the unobservable revenue y_u to the borrower, supplemented by a portion of observable revenue such that the borrower's total expected revenue equals $\hat{W}(n)$, i.e., setting $\overline{x}(y_z) = \tau y_z$, with $\mathbb{E}[y - \overline{x}(y_z)] = \mathbb{E}[y_u] + (1-\tau)\mathbb{E}[y_z] = \hat{W}(n)$, and $D(y_z) = \emptyset$.

However, this is not optimal, since the lender can always reduce α to the point such that $\mathbb{E}[y_u] = (1-\alpha)R_K = \hat{W}(n)$. In this case, all the unobservable revenue are kept by the borrower, while all the observable revenue are transferred to the lender (i.e., $\overline{x}(y_z) = y_z$ and $D(y_z) = \emptyset$). Thus, the optimal level of α should be low enough such that $\mathbb{E}[y_u] = (1-\alpha)R_K \geq \hat{W}(n)$ holds. This indicates that

$$\overline{x}(y_z) \geq y_z$$

should hold for all the realizations of observable revenue y_z.[A2]

In what follows, we consider the optimal repayment rules and ex-post verification region.

First, we show that given the realized y_z, the optimal verification region

[A2]Since $\mathbb{E}[y_u] \geq \hat{W}(n)$, if there exists some states y_z' such that $\overline{x}(y_z') < y_z'$, then there must exist some other states y_z'' where $\overline{x}(y_z'') > y_z''$. Otherwise, the borrower's expected revenue would be strictly higher than $\hat{W}(n)$ (i.e.,over-compensation), that is, $\mathbb{E}[y_u + y_z - x(y, y_z)] > \mathbb{E}[y_u] \geq \hat{W}_n$. Then, it is optimal to raise $\overline{x}(y_z')$ to y_z' while reducing y_z''.

has the property that[A3]

$$D(y_z) = [0, y^T(y_z)], \quad with \quad y^T(y_z) \geq y_z.$$

Suppose, to the contrary, that for a given y_z, the verification region is given as $D(y_z) = [0, y^T(y_z)] \cup \{y'\}$, where $y' > y^T(y_z)$ is an isolated point. Since the payment $\overline{x}(y_z)$ is feasible at $y = y^T(y_z)$, it must also be feasible when $y = y' > y^T(y_z)$. Thus, it is optimal to take y' out of the monitoring region, while simultaneously reducing $\overline{x}(y_z)$ such that the borrower's expected revenue remains at $\hat{W}(n)$ (see (A.2)). By doing so, the lender can get a higher expected revenue, as the verification region $D(y_z)$ becomes now smaller (see (A.1)).

Similarly, suppose that the verification region is given by the punctured interval $D(y_z) = [0, y') \cup (y', y^T(y_z)]$. Notice that since the transfer $\overline{x}(y_z)$ is feasible at $y = y'$, it must also be feasible for any $y \in (y', y^T(y_z)]$. Thus, it is optimal to take the interval $(y', y^T(y_z)]$ out of the monitoring region, while simultaneously reducing $\overline{x}(y_z)$ such that the borrower's expected revenue remains at $\hat{W}(n)$ (see (A.2)). By doing so, the lender can achieve a higher expected revenue, because the verification region $D(y_z)$ becomes now smaller (see (A.1)).

Second, for any given y_z, the optimal repayment rules have the following properties:

$$\overline{x}(y_z) = y^T(y_z) \quad and \quad \underline{x}(y, y_z) = y.$$

Note that $\overline{x}(y_z)$ cannot be larger than $y^T(y_z)$. Otherwise, the feasibility constraint (A.4) will be violated for some $y \in (y^T(y_z), \infty)$. Note also that it is not optimal for $\overline{x}(y_z)$ to be less than $y^T(y_z)$. In such cases, one can narrow the verification region by choosing a lower threshold, $\tilde{y}^T(y_z) = \overline{x}(y_z)$, and then reducing $\overline{x}(y_z)$ such that the borrower's expected revenue remains at $\hat{W}(n)$ (see (A.2)). As a result, the lender can get a higher expected revenue, as the verification region $D(y_z)$ becomes smaller (see (A.1)).

Also, suppose $\underline{x}(y', y_z) < y'$ for some $y' \in [0, y^T(y_z)]$, it is optimal to increase $\underline{x}(y', y_z)$ to y', and then reduce $\overline{x}(y_z)$ and $y^T(y_z)$ to keep the borrower's expected revenue remains at $\hat{W}(n)$ (see (A.2)). This narrows the verification region $D(y_z)$, allowing the lender to get a higher expected return (see (A.1)).

[A3] Notice that the threshold, $y^T(y_z)$, should be no less than y_z, since y_z is the lower bound of y (see equation (1)).

A1.2 Derivation of function $\Omega(u^T(z) - \kappa z)$

Note first that

$$\Gamma'(u^T(z)|z) = e^{u^T(z)}[1 - F_{u|z}(u^T(z)|z)] \quad \textit{and} \quad \Theta'(u^T(z)|z) = \mu e^{u^T(z)} f_{u|z}(u^T(z)|z),$$

where $f(u|z)$ and $F(u|z)$ are respectively the p.d.f. and c.d.f. of u conditional on z. Note also that

$$u|z \sim \mathcal{N}(\kappa z + \varsigma, \sigma^2_{u|z}),$$

as mentioned in Section 2.1. Thus,

$$f_{u|z}(u|z) = \tilde{f}(u - \kappa z) \quad \textit{and} \quad F_{u|z}(u|z) = \tilde{F}(u - \kappa z), \tag{A.5}$$

where $\tilde{f}(\cdot)$ and $\tilde{F}(\cdot)$ are respectively the p.d.f. and c.d.f. of $\mathcal{N}(\varsigma, \sigma^2_{u|z})$. Thus, we have

$$\frac{\Theta'(u^T(z)|z)}{\Gamma'(u^T(z)|z)} = \mu \cdot \frac{\tilde{f}(u - \kappa z)}{1 - \tilde{F}(u - \kappa z)} = \mu \cdot \Omega(u^T(z) - \kappa z).$$

A1.3 Properties of function $\Omega(u^T(z) - \kappa z)$

Note that function $\Omega(a)$ is given as

$$\Omega(a) = \frac{\tilde{f}(a)}{1 - \tilde{F}(a)},$$

where $\tilde{f}(\cdot)$ and $\tilde{F}(\cdot)$ are respectively the p.d.f. and c.d.f. of $\mathcal{N}(\varsigma, \sigma^2_{u|z})$.

Note also that

$$\Omega'(a) = \frac{\tilde{f}'(a)(1 - \tilde{F}(a)) + \tilde{f}^2(a)}{(1 - \tilde{F}(a))^2}.$$

Since $\tilde{f}'(a) = -[(a - \varsigma)/\sigma^2_{u|z}] \cdot \tilde{f}(a)$, thus $\Omega'(a)$ can be re-written as

$$\Omega'(a) = \frac{\tilde{f}(a)}{(1 - \tilde{F}(a))^2} \cdot \chi(a),$$

where

$$\chi(a) \equiv \tilde{f}(a) - \frac{a - \varsigma}{\sigma^2_{u|z}} \cdot (1 - \tilde{F}(a)).$$

It can be shown that $\chi'(a) = -(1 - \tilde{F}(a))/\sigma^2_{u|z} < 0$ and that $\chi(a)$ goes to zero when $a \to \infty$. Thus, $\chi(a)$ is always positive for any $a \in (-\infty, \infty)$. It follows that $\Omega'(a)$ is always positive.

A1.4 Derivation of $\hat{\Gamma}(u^*)$ and $\hat{\Theta}(u^*)$

Note first that by using (15) and (A.5), we have

$$\begin{aligned}
\Gamma(u^T(z)|z) = \Gamma(u^* + \kappa z|z) &= e^{u^*+\kappa z}[1 - F(u^* + \kappa z|z)] + \int_{-\infty}^{u^*+\kappa z} e^u f(u|z)du, \\
&= e^{u^*+\kappa z}[1 - \tilde{F}(u^*)] + e^{\kappa z} \int_{-\infty}^{u^*} e^u \tilde{f}(u)du, \\
&= e^{u^*+\kappa z}[1 - \tilde{F}(u^*)] + e^{\frac{\sigma_{s|z}^2}{2}+\kappa z} \cdot \tilde{F}(u^* - \sigma_{u|z}^2), \\
\Theta(u^T(z)|z) = \Theta(u^* + \kappa z|z) &= \mu \int_{-\infty}^{u^*+\kappa z} e^u f(u|z)du = \mu \cdot e^{\kappa z} \int_{-\infty}^{u^*} e^u \tilde{f}(u)du, \\
&= \mu \cdot e^{\frac{\sigma_{s|z}^2}{2}+\kappa z} \cdot \tilde{F}(u^* - \sigma_{u|z}^2).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\hat{\Gamma}(u^*) &\equiv \int_{-\infty}^{\infty} \Gamma(u^* + \kappa z|z)dG(z) = e^{u^*-\frac{\sigma_{s|z}^2}{2}}[1 - \tilde{F}(u^*)] + \tilde{F}(u^* - \sigma_{u|z}^2), \\
\hat{\Theta}(u^*) &\equiv \int_{-\infty}^{\infty} \Theta(u^* + \kappa z|z)dG(z) = \mu \cdot \tilde{F}(u^* - \sigma_{u|z}^2).
\end{aligned}$$

A1.5 Derivation of $\hat{\pi}(n)$

Note first that by rearranging the expression for $\hat{\pi}(n)$ in Problem 3, we have

$$\hat{\pi}(n) = \rho R_K - (1 - n)R + (1 - \alpha)(1 - \rho - \hat{\Theta}(u^*))R_K.$$

Note also that condition (18) implies

$$1 - \alpha = \frac{\rho R_K - (1 - n)R}{\rho R_K - [\hat{\Gamma}(u^*) - \hat{\Theta}(u^*)]R_K}.$$

By combing the above two equations, we have

$$\hat{\pi}(n) = [\rho R_K - (1 - n)R] \cdot \frac{1 - \hat{\Gamma}(u^*)}{\rho - [\hat{\Gamma}(u^*) - \hat{\Theta}(u^*)]}.$$

By using condition (17) in the above equation, we have

$$\hat{\pi}(n) = (1 + \lambda)[\rho R_K - (1 - n)R] = (1 + \lambda)R\left(n + \frac{\rho R_K - R}{R}\right).$$

A1.6 Derivation of market volatility σ_m^2

Note first that market return r_m, defined by (25), can be re-written as

$$r_m = \frac{\int_{i \in \mathcal{B}(s)} x_i}{\int_{i \in \mathcal{B}(s)} 1 - n_i},$$

where

$$x_i = \left[\alpha_i \cdot \rho \cdot e^{z_i} + (1 - \alpha_i) \cdot e^{u^* + \kappa z_i}\right] \cdot R_K.$$

Remember that $z_i = s + \varepsilon_{z,i}$, where the idiosyncratic shock $\varepsilon_{z,i}$ is i.i.d. across borrowers and is independent of the borrowers' choices of α_i and n_i. Thus, given the aggregate state s, the market average of repayment x_i (conditional on non-liquidation), denoted by $x_m(s)$, is given as

$$x_m(s) = \left[\alpha_m \cdot \rho \cdot \phi_1(s) + (1 - \alpha_m) \cdot e^{u^*} \cdot \phi_\kappa(s)\right] \cdot R_K, \tag{A.6}$$

where α_m is the simple average of α_i, and

$$\phi_1(s) = e^s \cdot \mathbb{E}_\varepsilon(e^{\varepsilon_{z,i}} | i \in \mathcal{B}(s)) \quad \textit{and} \quad \phi_\kappa(s) = e^{\kappa s} \cdot \mathbb{E}_\varepsilon(e^{\kappa \varepsilon_{z,i}} | i \in \mathcal{B}(s)). \tag{A.7}$$

Then, the ex-ante market volatility, σ_m^2, defined by (26), can be re-written as

$$\sigma_m^2 = \frac{\mathbb{E}_s(x_m(s)^2) - \mathbb{E}_s(x_m(s))^2}{(1 - n_m)^2},$$

where n_m is the simple average of n_i, and

$$\mathbb{E}_s(x_m(s)) = \left\{\alpha_m \cdot \rho \cdot \mathbb{E}_s[\phi_1(s)] + (1 - \alpha_m) \cdot e^{u^*} \cdot \mathbb{E}_s[\phi_\kappa(s)]\right\} \cdot R_K, \tag{A.8}$$

$$\mathbb{E}_s(x_m(s)^2) = \mathbb{E}_s\left\{\left[\alpha_m \cdot \rho \cdot \phi_1(s) + (1 - \alpha_m) \cdot e^{u^*} \cdot \phi_\kappa(s)\right]^2\right\} \cdot R_K^2. \tag{A.9}$$

To compute the value of σ_m^2, we first derive analytically $\mathbb{E}_\varepsilon(e^{\varepsilon_{z,i}} | i \in \mathcal{B}(s))$ and $\mathbb{E}_\varepsilon(e^{\kappa \varepsilon_{z,i}} | i \in \mathcal{B}(s))$ in expressions (A.7). Recall that liquidation does not happen ($i \in \mathcal{B}(s)$) when $u_i > u^* + \kappa z_i$. Since $z_i = s + \varepsilon_{z,i}$ and $u_i = s + \varepsilon_{u,i}$, thus the condition for non-liquidation can be re-written as

$$\varepsilon_{u,i} > u^* - (1 - \kappa)s + \kappa \varepsilon_{z,i}.$$

Thus, the probability density function of $\varepsilon_{z,i}$ conditional on non-liquidation is given as

$$f_c(\varepsilon_z | i \in \mathcal{B}(s)) = \frac{f_{\varepsilon_z}(\varepsilon_z) \cdot [1 - F_{\varepsilon_u}(u^* - (1 - \kappa)s + \kappa \varepsilon_z)]}{1 - F_{\varepsilon_u - \kappa \varepsilon_z}(u^* - (1 - \kappa)s)}$$

By using the above probability density function, it can be shown that

$$\mathbb{E}_{\varepsilon}(e^{\varepsilon_{z,i}}|i \in \mathcal{B}(s)) = \frac{1 - F_{\varepsilon_u - \kappa\varepsilon_z}(u^* - (1-\kappa)s + \kappa\sigma_{\varepsilon}^2)}{1 - F_{\varepsilon_u - \kappa\varepsilon_z}(u^* - (1-\kappa)s)}.$$

Similarly, we can show that

$$\mathbb{E}_{\varepsilon}(e^{\kappa\varepsilon_{z,i}}|i \in \mathcal{B}(s)) = e^{\kappa(1-\kappa)\varsigma} \cdot \frac{1 - F_{\varepsilon_u - \kappa\varepsilon_z}(u^* - (1-\kappa)s + \kappa^2\sigma_{\varepsilon}^2)}{1 - F_{\varepsilon_u - \kappa\varepsilon_z}(u^* - (1-\kappa)s)}.$$

Thus, by using the above results in expressions (A.7), we obtain analytical expressions for $\phi_1(s)$ and $\phi_\kappa(s)$. After that, the values of $\mathbb{E}_s[\phi_1(s)]$, $\mathbb{E}_s[\phi_1(s)^2]$, $\mathbb{E}_s[\phi_\kappa(s)]$ and $\mathbb{E}_s[\phi_\kappa(s)^2]$ in expressions (A.8) and (A.9) are computed numerically.

A1.7 Computation of the impulse responses

We consider impulse responses to a negative shock to the aggregate productivity, s, of size $2\sigma_s$. Remember that all the shocks in the model have no persistence, and the only aggregate state is the wealth distribution of the borrowers, $\Phi(w)$. We assume that the economy is initially at the steady state, and the shock hits the economy in period 1. The impulse responses are computed using the Monte Carlo techniques developed by Koop et al. (1996). The impulse response function is defined as:

$$\mathbf{GI}_X(t|s_1, \Phi_0(w)) = \mathbb{E}(\mathbf{X}_t|s_1, \Phi_0(w)) - \mathbb{E}(\mathbf{X}_t|\Phi_0(w)), \quad for \quad t = 1, 2, ..., T,$$

where T is the horizon of the impulse responses, $\mathbb{E}(\mathbf{X}_t|\Phi_0(w))$ is the expectation of $\mathbf{X}_t$ conditional on the initial state, $\Phi_0(w)$, and $\mathbb{E}(\mathbf{X}_t|s_1, \Phi_0(w))$ is the expectation of $\mathbf{X}_t$ conditional on not only $\Phi_0(w)$ but also s_1. Recall that s_1 is the aggregate productivity shock in period 1.

We set the number of replications to R, and draw independently innovations from normal distributions. The initial state, $\Phi_0(w)$, corresponds to the wealth distribution at the non-stochastic steady state. To compute $\mathbb{E}(\mathbf{X}_t|\Phi_0(w))$ and $\mathbb{E}(\mathbf{X}_t|s_1, \Phi_0(w))$ numerically, we conduct the following two simulations.

First, we simulate the model without the shock. For each replication r, given the initial wealth distribution, $\Phi_0(w)$, we simulate the model using randomly drawn innovations and obtain $\mathbf{X}_t^{(r)}$ for $t = 1, 2, ..., T$ and $r = 1, 2, ..., R$.

Second, we simulate the model with the shock. For each replication r, given the initial wealth distribution, $\Phi_0(w)$, we simulate the model using randomly drawn innovations with the aggregate shock in period 1 being replaced by $s_1 = -2\sigma_s$. Then, we obtain $\check{\mathbf{X}}_t^{(r)}$ for $t = 1, 2, ..., T$ and $r = 1, 2, ..., R$.

The two expectations are approximated as the average across replications:

$$\hat{\mathbb{E}}(\mathbf{X}_t | s_1, \Phi_0(w)) = \frac{1}{R}\sum_{r=1}^{R} \check{\mathbf{X}}_t^{(r)} \quad \textit{and} \quad \hat{\mathbb{E}}(\mathbf{X}_t | \Phi_0(w)) = \frac{1}{R}\sum_{r=1}^{R} \mathbf{X}_t^{(r)}.$$

Then, the approximated impulse response in period t is given by

$$\mathbf{GI}_X(t | s_1, \Phi_0(w)) = \frac{1}{R}\sum_{r=1}^{R} \check{\mathbf{X}}_t^{(r)} - \frac{1}{R}\sum_{r=1}^{R} \mathbf{X}_t^{(r)}.$$

Finally, we compute the long-run unconditional means of the variables, denoted by $\mathbb{E}_0(\mathbf{X})$, by simulating the model for 12,000 periods, while using the wealth distribution at the non-stochastic steady state as the initial state. To obtain $\hat{\mathbb{E}}_0(\mathbf{X})$, we drop the first 2000 periods and take time average of the simulated values. Then, we divide the approximated impulse responses by corresponding long-run unconditional means:

$$\frac{\mathbf{GI}_X(t | s_1, \Phi_0(w))}{\hat{\mathbb{E}}_0(\mathbf{X})}.$$

A2 Calibration and numerical exercises

A2.1 Correlations

Table A1: Correlations

	I. Cross-sectional standard deviation of stock returns		II. Volatility of market return	
	Data (1)	Model (2)	Data (3)	Model (4)
TFP_{t-1}	-0.768***	-0.611***	-0.717***	-0.841***
	(0.102)	(0.003)	(0.091)	(0.002)
R^2	0.484	0.373	0.503	0.708
Observations	63	-	63	-

Notes: The basket of selected stocks are consistent with Bloom (2009), and market volatility is measured using the VXO index. The actual tightness of the economy-wide financial condition is measured using the Chicago Fed National Financial Condition Index. The actual TFP shocks (to business sector) are estimated by Fernald (2014) and downloaded from the Federal Reserve Bank of San Fransisco, which are available only at quarterly frequency. Thus, the actual correlations between $\sigma_{p,t}$ ($\sigma_{m,t}$) and TFP shocks are estimated using quarterly data. All variables are normalized to have a standard deviation of one. Standard errors are reported in the brackets below.

A2.2 Robustness

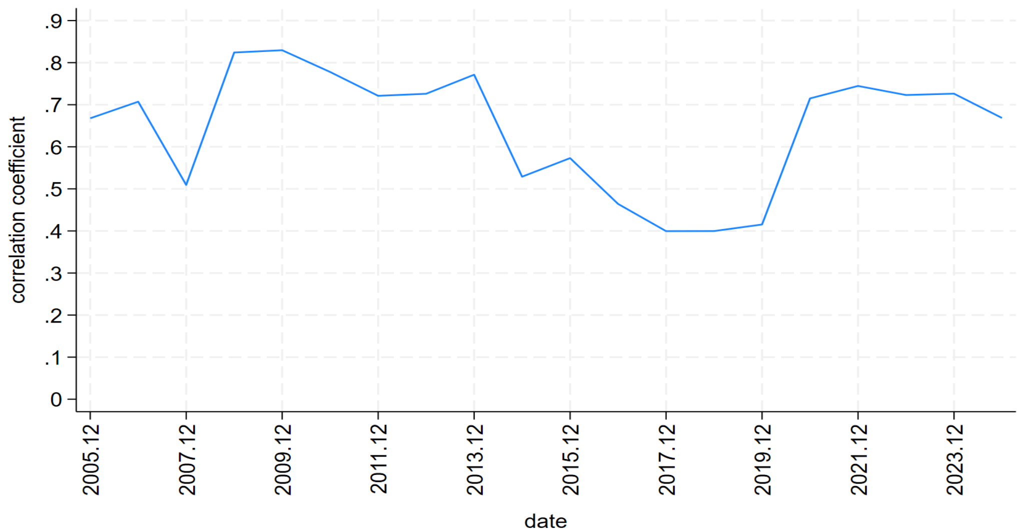

Figure A1: Correlation coefficient between the actual cross-sectional standard deviation of individual stock returns and the VXO index

Notes: The graph shows the correlation coefficient between the cross-sectional standard deviation of actual individual stock returns and the VXO index over a 60 month rolling window, with step size being set to 12 month.

Table A2: Robustness of model results to changes in value of ρ

	Data	Model				
		The cost of ex-ante monitoring, $1-\rho=$				
		Baseline	Lower cost		Higher cost	
		0.024	0.012	0.018	0.030	0.036
	(1)	(2)	(3)	(4)	(5)	(6)
$corr(\sigma_{p,t}, \sigma_{m,t})$	0.719	0.720	0.747	0.734	0.705	0.687
$corr(\sigma_{p,t}, finc_t)$	0.795	0.719	0.745	0.732	0.704	0.686
$corr(\sigma_{m,t}, finc_t)$	0.842	0.999	0.999	0.999	0.999	0.999
$corr(\sigma_{p,t}, tfp_{t-1})$	-0.768	-0.611	-0.626	-0.620	-0.601	-0.588
$corr(\sigma_{m,t}, tfp_{t-1})$	-0.717	-0.841	-0.834	-0.839	-0.845	-0.847
$\sigma_{p,ss}$	0.081	0.067	0.065	0.066	0.068	0.070
$\sigma_{m,ss}$	0.064	0.067	0.064	0.065	0.068	0.070

Notes: This table shows the key model moments for different values of ρ. The first to the fifth rows show the pairwise correlations between cross-sectional standard deviation of individual stock returns $\sigma_{p,t}$, volatility of market return $\sigma_{m,t}$, the economy-wide financial condition $finc_t$, and aggregate TFP shocks. The sixth to the seventh rows show the mean of $\sigma_{p,t}$ and $\sigma_{m,t}$ at the stochastic steady state. The actual economy-wide financial condition is proxied by the NFCI index, while in the model it is represented by the spread between aggregate investment return and funding cost ($R_K - R$). The actual market volatility $\sigma_{m,t}$ is proxied using the VXO index (de-annualized). The actual TFP shocks (to business sector) are estimated by Fernald (2014) and downloaded from the Federal Reserve Bank of San Fransisco, which are available only at quarterly frequency. Thus, the actual correlations between $\sigma_{p,t}$ ($\sigma_{m,t}$) and TFP shocks are computed using quarterly data.

Table A3: Robustness of model results to changes in values of σ_s and σ_ε

	Data	Model				
		The std. of aggregate and idiosyncratic shocks, $\sigma_s = \sigma_\varepsilon =$				
		Baseline	Lower risks		Higher risks	
		0.1	0.05	0.075	0.125	0.15
	(1)	(2)	(3)	(4)	(5)	(6)
$corr(\sigma_{p,t}, \sigma_{m,t})$	0.719	0.720	0.869	0.803	0.635	0.556
$corr(\sigma_{p,t}, finc_t)$	0.795	0.719	0.869	0.802	0.633	0.554
$corr(\sigma_{m,t}, finc_t)$	0.842	0.999	0.999	0.999	0.999	0.998
$corr(\sigma_{p,t}, tfp_{t-1})$	-0.768	-0.611	-0.810	-0.712	-0.520	-0.442
$corr(\sigma_{m,t}, tfp_{t-1})$	-0.717	-0.841	-0.918	-0.877	-0.815	-0.794
$\sigma_{p,ss}$	0.081	0.067	0.029	0.047	0.088	0.110
$\sigma_{m,ss}$	0.064	0.067	0.029	0.047	0.088	0.111

Notes: This table shows the key model moments for different values of σ_s and σ_ε. The first to the fifth rows show the pairwise correlations between cross-sectional dispersion of individual stock returns $\sigma_{p,t}$, volatility of market return $\sigma_{m,t}$, the economy-wide financial condition $finc_t$, and aggregate TFP shocks. The sixth to the seventh rows show the mean of $\sigma_{p,t}$ and $\sigma_{m,t}$ at the stochastic steady state. The actual economy-wide financial condition is proxied by the NFCI index, while in the model it is represented by the spread between aggregate investment return and funding cost ($R_K - R$). The actual market volatility $\sigma_{m,t}$ is proxied using the VXO index (de-annualized). The actual TFP shocks (to business sector) are estimated by Fernald (2014) and downloaded from the Federal Reserve Bank of San Fransisco, which are available only at quarterly frequency. Thus, the actual correlations between $\sigma_{p,t}$ ($\sigma_{m,t}$) and TFP shocks are computed using quarterly data.

A2.3 Impulse responses

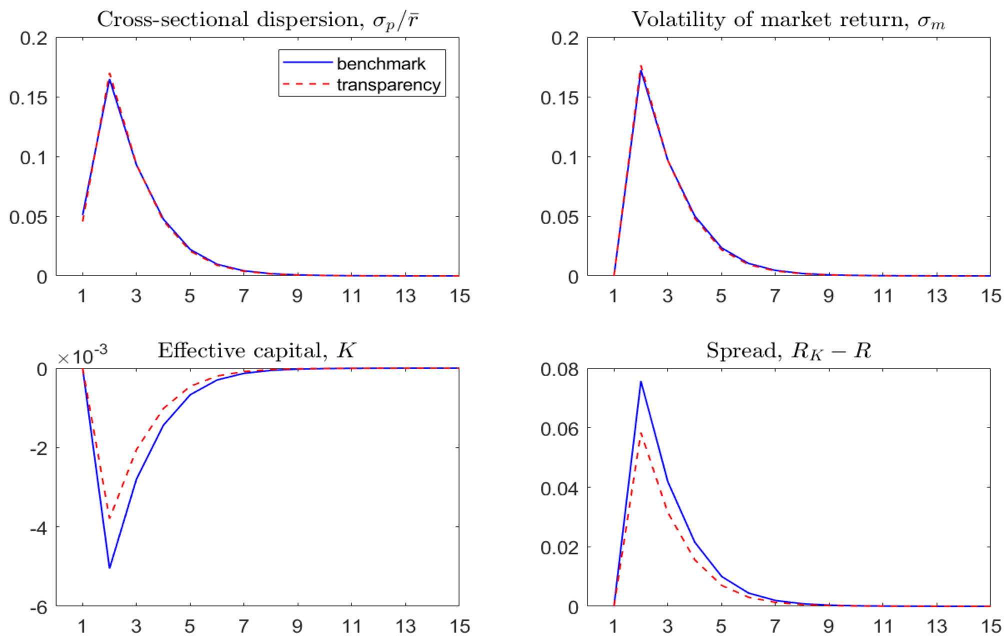

Figure A2: Responses to a negative aggregate TFP shock under different levels of transparency

Notes: The solid-blue lines show responses of the benchmark economy. The dashed-red lines show responses of the alternative economy with a higher level of transparency (in which the monitoring costs, μ and $1-\rho$, are 25% lower compared to those in the benchmark model). All variables are normalized by corresponding long-run unconditional means. Time horizon in months.

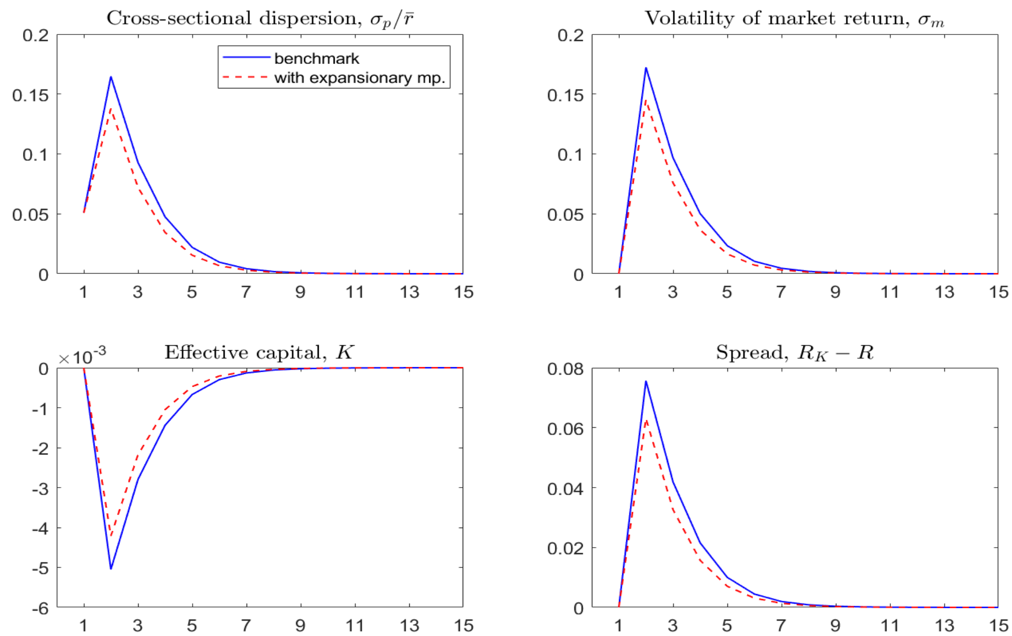

Figure A3: Responses to a negative aggregate TFP shock with and without monetary policy stimulus

Notes: The solid-blue lines show responses of the benchmark economy without monetary policy. The dashed-red lines show responses of an alternative economy in which monetary policy reacts to the negative shock by reducing real interest rate by 1% in the period after the shock. All variables are normalized by corresponding long-run unconditional means. Time horizon in months.

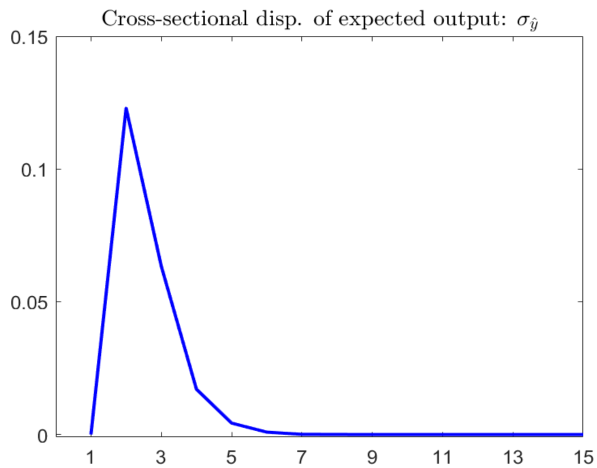

Figure A4: Responses of cross-sectional dispersion in expected output to a negative aggregate TFP shock

Notes: Variables are normalized by corresponding long-run unconditional means. Time horizon in months.

A3 Extensions

A3.1 Unobservable α

We now consider the case where setting the intensity of ex-ante monitoring, α, is an action to be taken by the borrower and is unobservable to the lender. Consequently, α is non-contractible. As will be shown, the lender needs to provide additional incentives to prevent the borrower from deviating from the prescribed α, thereby inducing additional costs associated with ex-ante monitoring. However, the model results remain qualitatively the same as in the baseline model: in equilibrium the intensity of ex-ante monitoring, α, is decreasing in the borrower's net worth.

We start by showing why the borrower has an incentive to deviate from the prescribed α.

For illustrative purpose, we consider a simplified version of the model in which the macro component of productivity shocks, s, is set to zero. Thus, shocks u and z are now independent from each other, indicating that $\kappa = 0$ (see (3) and (4)). Under this specification, the optimal repayment rules and verification region are given as (see Proposition 1)

$$D(y_z) = [0, y^T(y_z)], \quad \textit{where} \quad y^T(y_z) = y_z + y_u^T, \tag{A.10}$$

and

$$x(y, y_z) = \begin{cases} y^T(y_z), & \textit{if } y \notin D(y_z), \\ y, & \textit{otherwise}. \end{cases} \tag{A.11}$$

Therefore, the borrower's expected revenue is

$$\mathbb{E}[y - x(y, y_z)] = \int_{y_u^T}^{\infty} (y_u - y_u^T)\, dF_{y_u}(y_u; \alpha),$$

while the expected transfer to the lender is

$$\mathbb{E}[x(y, y_z)] = \int_0^{\infty} y_z\, dF_{y_z}(y_z; \alpha) + y_u^T \cdot \left(1 - F_{y_u}(y_u; \alpha)\right) + \int_0^{y_u^T} y_u\, dF_{y_u}(y_u; \alpha).$$

As can be seen from the above expressions, under the current contract all the observable revue y_z are transferred to the lender, while the borrower retains the portion of y_u that exceeds the threshold y_u^T. By lowering α, the probability mass function of y_u is shifted to the right, while the probability mass function of y_z is shifted to the left. This raises the borrower's expected revenue, while reduces the expected transfer to the lender.

Performance-based hurdle. We show that the lender can use a simple performance-based hurdle to incentivize the borrower to adhere to the prescribed α. As explained before, the borrower has an incentive to lower α, which shifts the probability mass function of y_z to the left, lowering the expected value of y_z. To penalize such deviation, the lender can implement a performance-based hurdle: the borrower receives zero revenue when the realized observable performance y_z falls below a threshold y_z^T. Admittedly, there may exist alternative ways to incentivize the borrower. However, we restrict our attention to this type of performance hurdle because it is simple to implement and widely observed in the practice.

With the performance-based hurdle, the repayment rules and verification region are changed from (A.10) and (A.11) to the following:

$$D(y_z) = \begin{cases} [0, y^T(y_z)], & \textit{if } y_z > y_z^T, \\ [0, y], & \textit{if } y_z \le y_z^T, \end{cases} \tag{A.12}$$

where $y^T(y_z) = y_z + y_u^T$, and

$$x(y, y_z) = \begin{cases} y^T(y_z), & \textit{if } y_z > y_z^T \textit{ and } y \notin D(y_z), \\ y, & \textit{otherwise.} \end{cases} \tag{A.13}$$

Figure A5 below shows the borrower's expected revenue (left panel) and the expected transfer to the lender (right panel), conditional on the realization of y_z. As shown in the figure, when the observed performance y_z is below the threshold y_z^T, the borrower's revenue is zero. Once the realized y_z exceeds y_z^T, the borrower's expected revenue jumps to a strictly positive value. When the borrower lowers α, the probability mass function of y_z is shifted to the left, and the observable performance y_z becomes more likely to fall below y_z^T. This discourages the borrower from deviating from the prescribed α.

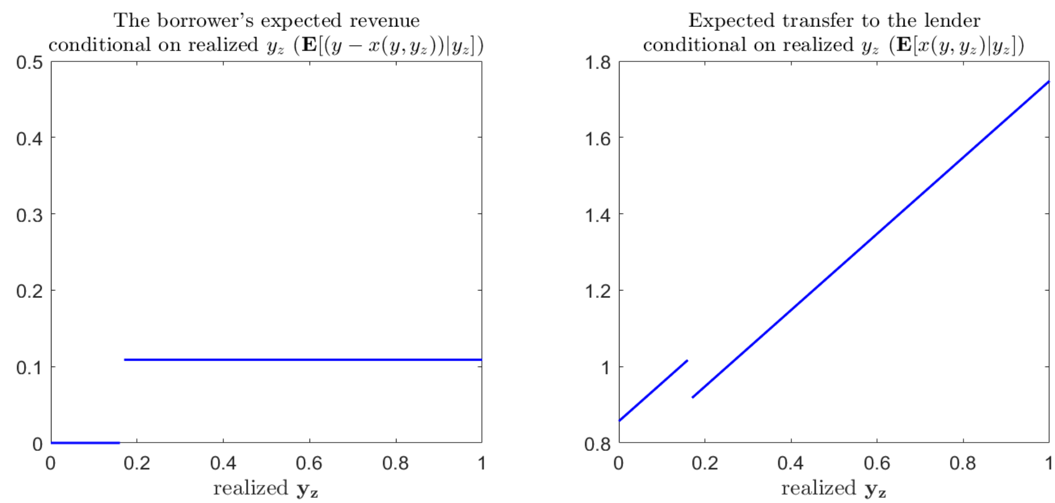

Figure A5: Performance-based hurdle

Optimal α. We now consider how the intensity of ex-ante monitoring, α, is determined. Note that with the performance-based hurdle, the lender's optimization problem is given as follows:

$$\hat{\Pi}(n) = \max_{\alpha, y_z^T, y_u^T} \hat{y}_z(\alpha) + F_{y_z}(y_z^T; \alpha) \cdot (1-\mu) \cdot \hat{y}_u(\alpha) \tag{A.14}$$
$$+ \Big(1 - F_{y_z}(y_z^T; \alpha)\Big) \cdot \left[y_u^T \cdot \Big(1 - F_{y_u}(y_u^T)\Big) + (1-\mu) \int_0^{y_u^T} y_u \, dF_{y_u}(y_u; \alpha) \right]$$

subject to the promised return constraint

$$\tilde{W}(\alpha, y_z^T, y_u^T) \equiv \Big(1 - F_{y_z}(y_z^T; \alpha)\Big) \cdot \int_{y_u^T}^{\infty} (y_u - y_u^T) \, dF_{y_u}(y_u; \alpha) \geq \hat{W}(n). \tag{A.15}$$

and the incentive-compatible constraint

$$\alpha \textit{ maximizes } \quad \tilde{W}(\alpha, y_z^T, y_u^T),$$

which implies

$$\frac{\partial \tilde{W}(\alpha, y_z^T, y_u^T)}{\partial \alpha} = -\frac{\partial F_{y_z}(y_z^T; \alpha)}{\partial \alpha} \cdot \int_{y_u^T}^{\infty} (y_u - y_u^T) \, dF_{y_u}(y_u; \alpha)$$
$$- \Big(1 - F_{y_z}(y_z^T; \alpha)\Big) \cdot \left(\frac{1}{1-\alpha}\right) \cdot \int_{y_u^T}^{\infty} y_u \, dF_{y_u}(y_u; \alpha) = 0. \tag{A.16}$$

Here, function $\tilde{W}(\alpha, y_z^T, y_u^T)$ represents the borrower's expected revenue. Since α is non-contractible, it must maximize $\tilde{W}(\alpha, y_z^T, y_u^T)$ to satisfy the incentive compatibility constraint; Otherwise, the borrower will deviate from the prescribed α. Condition (A.16) is the first-order condition, which is a necessary condition for α to maximize $\tilde{W}(\alpha, y_z^T, y_u^T)$ given y_z^T and y_u^T.

We solve for the optimal α numerically. Specifically, for a given value of α, we solve for y_z^T and y_u^T using (A.15) and (A.16), and then compute the lender's expected revenue.[A4] We search over the interval $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ to find the optimal α^* that maximizes the lender's expected revenue.

The results are plotted in Figure A6 below. The left panel of the figure shows the baseline case where the borrower's required revenue $\hat{W}(n)$ is set to 0.1. In this case, the lender's expected return is maximized when $\alpha^* = 0.187$. We then consider a scenario where the borrower invests less net worth into the project, such that $\hat{W}(n)$ decreases to 0.095. As a result, the optimal α^* rises to 0.217. These results are qualitatively consistent with the baseline model: a lower level of net worth is associated with a higher level of ex-ante monitoring.

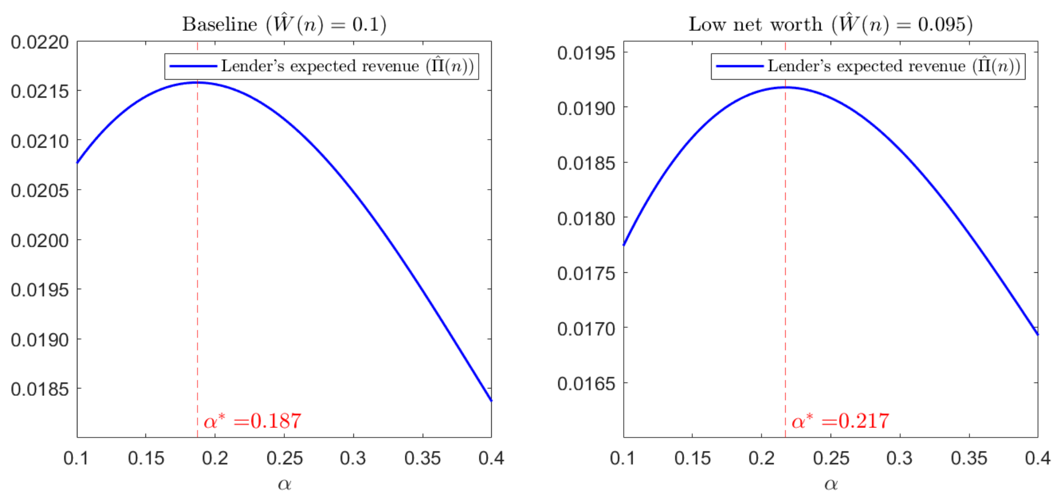

Figure A6: Optimal α^*

To confirm the global optimality of α^* from the borrower's perspective, we compute the borrower' expected revenue, $\tilde{W}(\alpha, y_z^T, y_u^T)$, for different values of α, while fixing y_z^T and y_u^T at their contractual levels. The results

[A4] Note that given α, y_z^T and y_u^T are implicitly defined by (A.15) and (A.16).

are shown in Figure A7. As can be seen, the borrower's expected revenue is indeed maximized at $\alpha = \alpha^*$, and thus the borrower has no incentive to deviate from the prescribed α^*.

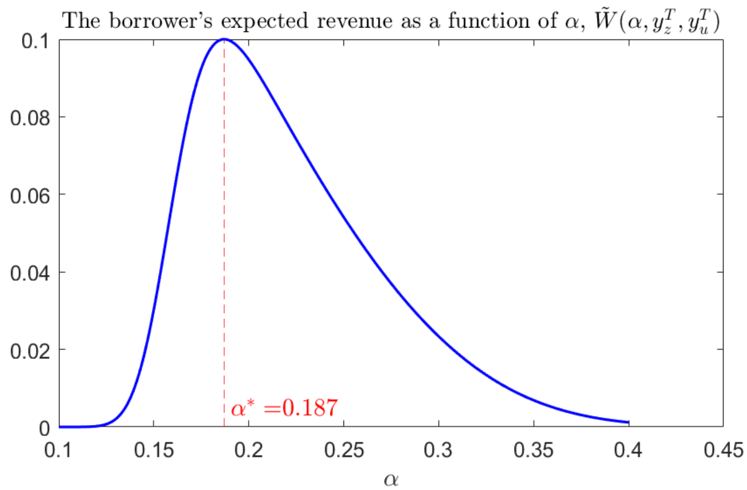

Figure A7: Global optimality of α^* from the borrower's perspective

A3.2 Contracts offered by the borrower

In this subsection, we show that given the bargaining powers of the lender and the borrower, the optimal contract is invariant to who proposes the contract.

Note that if the contract is offered by the borrower, Problem 3 is changed to

$$\hat{W}(n) = \max_{\alpha\in[0,1],\{u^T(z)\}} \hat{y}_u(\alpha)\cdot\left(1-\int_{-\infty}^{\infty}\Gamma(u^T(z)|z)\,dF_z(z)\right),$$

subject to

$$\hat{\Pi}(n) = \hat{y}_z(\alpha) + \hat{y}_u(\alpha)\cdot\int_{-\infty}^{\infty}\left[\Gamma(u^T(z)|z) - \Theta(u^T(z)|z)\right]dF_z(z) - (1-n)\cdot R = 0.$$

Note that by combing the above two expressions, we have

$$\hat{W}(n) + \hat{\Pi}(n) = \max_{\alpha\in[0,1],\{u^T(z)\}} \hat{y}(\alpha) - \hat{y}_u(\alpha)\cdot\int_{-\infty}^{\infty}\Theta(u^T(z)|z)\,dF_z(z) - (1-n)\cdot R,$$

Note also that by combing the promised return constraint and the objective function in Problem 3, we obtain an expression identical to the one derived above. This indicates that no matter which party proposes the contract, the optimal contract should maximize the total surplus, $\hat{W}(n)+\hat{\Pi}(n)$. Under the assumption that lenders are perfectly competitive, the lender's expected net revenue should always be zero, $\hat{\Pi}(n) = 0$, indicating that the borrower always captures all the surplus. Thus, the borrower's expected revenue, $\hat{W}(n)$, remains the same no matter who proposes the contract.