# Asset Pricing with Heterogeneous Investors*

François Geerolf
UCLA

June 2019
**[Last Version]**

**Abstract**

I propose a competitive asset pricing model based on the scarcity and heterogeneity of lenders. Expected excess returns on a given investment do not just depend on the investment itself and the identity of the investor, but on aggregate factors such as the whole distribution of investors' wealth and borrowers' wealth, investors' heterogeneity and borrowers' heterogeneity. The model implies that expected returns have a common factor which cannot be traced to characteristics of asset, or the investor holding this asset, unlike in both conventional asset pricing and intermediary-based asset pricing. These asset pricing implications are obtained by microfounding lenders' technologies explicitly, in line with textbook models of corporate finance and financial frictions, and allowing heterogeneity among them; it is conceptually different from existing theories, in that it does not rely on risk aversion or liquidity. Applications of the model to syndicated loans, private equity, and venture capital, where financial frictions are important are discussed: the model explains some empirical facts which have previously been documented on these markets.

**Keywords:** Asset Pricing, Corporate Finance, Financial Frictions
**JEL classification:** G1, G10, G12.

*E-mail: fgeerolf@econ.ucla.edu. I thank Andy Abel, Andy Atkeson, Bhagwan Chowdhry, John Cochrane, Peter DeMarzo, Darrell Duffie, Andrea Eisfeldt, Mark Gertler, Simon Gilchrist, Lars Hansen, Arvind Krishnamurthy, Francis Longstaff, Timothy McQuade, Cyril Monnet, Pierre-Olivier Weill, and seminar participants at NYU Stern, UCLA, Stanford GSB, the University of Bern, the 2017 SED Meetings in Edinburgh, the 9th Sciences Po Summer Macro-Finance Conference, for useful comments. This is a revised version of a previous draft entitled "A Theory of Discount Rates".

# Introduction

Textbook finance theory is fundamentally based on investors' risk aversion, which creates a trade-off between return and risk. A basic prediction of asset pricing theory is that assets which pay more in states where investors have low consumption or wealth have higher expected returns on average, because of decreasing marginal utility. Despite its simplicity and theoretical appeal, this view of the return to capital has a hard time explaining some important empirical facts. Riskless arbitrage opportunities, or "fire sales", sometimes appear, which is puzzling for any theory based on a trade-off between risk and return, and capital moves only slowly to arbitrage them (Duffie (2010)). Average expected returns in the cross-section depend on "factors" other than aggregate consumption or any other measure of systematic risk (Fama and French (1992)): for example, distressed stocks have low prices and high expected returns (the "value" premium). According to Cochrane (2011, 2017), a major proponent of the risk-return approach to asset pricing, time series changes in expected returns are often still quite challenging to rationalize in many markets using aggregate consumption.

Intermediary-based asset pricing models, such as He and Krishnamurthy (2013) or Brunnermeier and Sannikov (2014), are also based on a risk-return trade-off, but applied to risk-averse financial institutions. In this theory, households are altogether unable to participate in financial markets, which explains the disconnect between the stochastic discount factor of the representative investor and asset prices. However, financial institutions are risk averse, and since they ultimately hold most of the capital stock, their stochastic discount factor should be used for asset pricing. This approach has has some empirical success: for example, the intermediary asset pricing factor of Adrian et al. (2014) has had considerable success in pricing the main cross-sectional anomalies of Fama and French (1992). He et al. (2017) show that corporate and sovereign bonds are related to the equity capital ratio of primary dealers. Similarly, credit spreads are more closely related to the distress of financial intermediaries than to consumption related variables, as shown by Muir (2017) and Krishnamurthy and Muir (2017). Applying the risk-return trade-off to financial institutions however leaves some theoretical and empirical questions unanswered. In particular, it is unclear in these models what financial institutions do. There is a stark segmentation between intermediaries who may participate in all markets, and households who participate in none. It is unclear what the microfoundations for financial institutions' risk aversion are, and why financial institutions become more risk averse, not less, when they approach bankruptcy.[1]

This paper offers an alternative view of intermediary-based asset pricing. Instead of

[1]Cochrane (2017) views this as a major theoretical difficulty for intermediary-based asset pricing models: "First, why do people get more risk averse as they approach bankruptcy, not less? (...) The usual concern is therefore that people and businesses near bankruptcy have incentives to take too much risk, not too little."

assuming stark market segmentation, the paper seeks to microfound what intermediaries do explicitly, in line with textbook models of corporate finance and financial frictions. The main advantage of such an approach is that it does not rely on intermediaries' risk aversion. Only some heterogeneity in financial intermediaries is needed: that may be reflecting heterogeneity in monitoring technologies, in second-best use values of the borrowers' assets, or of recovery value in bankruptcy. At the same time, the theory has several asset pricing implications in common with traditional intermediary-based asset pricing, and can therefore relate to much of the evidence supportive of this theory. It remains true that assets with a higher equilibrium covariance with intermediary leverage have higher expected returns on average, so that the theory is consistent with the evidence supportive of intermediary-based asset pricing. However, the mechanism does not rely on intermediaries' risk aversion, but simply on supply and demand for financial expertise, and the scarcity thereof. This difference is not just conceptual but also leads to distinctive empirical predictions: for example, apparently "riskless" arbitrage opportunities, or "fire sales", may open up when asset pricing is based on the scarcity of financial expertise. In other words, identical securities in the Arrow-Debreu sense can have different prices in equilibrium. This cannot happen with traditional intermediary-based asset pricing.

The basic argument follows directly from any textbook model of financing frictions, provided that there exists some heterogeneity on the side of investors. For instance, assume that a mass of identical financially constrained borrowers are bidding competitively for financing from both an expert investor, who has only limited financial resources, and an unsophisticated investor. Further assume that the expert investor can lend to borrowers, while requiring less "skin in the game" from them. This could arise from higher expertise, allowing him to reduce the scope for shirking, or from a better second-best use value of the asset, or a better ability to recover some value from the asset in bankruptcy, etc. In equilibrium, borrowing with lower skin in the game should be more expensive than borrowing with higher skin in the game. Otherwise all borrowers would prefer to borrow with lower skin in the game, and markets would not clear: indeed, provided that they have a productive technology, borrowers would ideally like to lever up as much as possible. Therefore, the price of assets that experts invest in should be lower than the price of assets that unsophisticated investor buy, everything else equal. Lending with lower skin in the game should thus earn higher expected returns. This means that assets with similar cash flows in all states of the world will have different prices, depending on how much skin in the game the corresponding investors require, a violation of no-arbitrage relations.

The reasoning is also very straightforward, and is only based on the supply and demand for different levels of expertise, taking the form of heterogenous pledgeability parameters, which also makes it quite general. In particular, it applies to many types of contracting structures: debt, equity, or more complex contingencies. The previous intuition can be generalized to a case where borrowers are heterogeneous as well. If entrepreneurs have heterogeneous productivities, then expected returns do not just make entrepreneurs indifferent between several types of contracts. Expected returns on high leverage ratio loans are higher, so that only more productive entrepreneurs want to borrow with higher leverage. Indeed, because of a complementarity between leverage and productivity, the competitive equilibrium has positive sorting between high productivity entrepreneurs and high expertise lenders. Expected returns correspond to walrasian prices which sustain this allocation.

To summarize, according to this corporate finance view of asset pricing, expected returns on assets may be determined by competition for scarce monitoring abilities, among financially constrained borrowers. Heterogenous and time varying expected returns correspond to the changing prices for different levels of financial expertise. In this theory, it does not matter how much curvature financial intermediaries' objective functions have. The response of financial intermediaries' risk aversion to shocks is also irrelevant, for determining whether expected returns increase when balance sheets are weak. In such a theory, expected returns increase when balance sheets are weak just because competition for monitoring abilities becomes more fierce. Using minimal ingredients, the model conveys the very simple idea that credit is more expensive (expected returns are higher), when credit supply is lower, or when demand for credit is higher. Other determinants of expected returns include borrowers' wealth, but also intermediaries' heterogeneity and borrowers' heterogeneity. All these comparative statics exercises would be puzzling from the point of view of traditional asset pricing theory.

Finally, potential applications of this asset pricing model to different market segments are discussed. A first market where monitoring is most salient is the private equity and venture capital market. I show that earlier evidence on these markets is very supportive of the theory. In particular, venture capital fund returns have been found to be very heterogeneous and persistent (Kaplan and Schoar (2005)). It has also been shown that better venture capital funds are able to negotiate better average returns ex-ante, by negotiating "better deal terms", in industry parlance. This is direct evidence supportive of the main message of this paper: asset prices can indeed reward scarce expertise. A second potential application of the model concerns mutual fund returns. If mutual funds do not just pick stocks but also vote in shareholder's meetings and are actively monitoring the companies they invest in (which can be more indirect, such as voting with activist hedge funds), then the gross of fees performance of mutual funds could also be coming from this return to financial expertise. However, the model would also predict that net of fees, no particular alpha should be obtained, which is also in line with the data. The third application concerns the public equity markets. I start from Adrian et al. (2014), who document the ability of intermediary asset pricing model to price equity portfolios (book to market, size, momentum) and ask what assumptions in the model may allow to microfound this correlation. I show that to the extent that small and distressed stocks are more likely to be held by experts, which could come from higher financial constraints or higher importance of monitoring when firms are "distressed", then indeed the model can rationalize why these firms have higher returns, and why their stock prices comove with intermediaries' leverage. Finally, again motivated by the evidence in favor of intermediary based asset pricing shown by He et al. (2017), I similarly discuss the valuation of corporate and sovereign bonds through the lens of the model. Again, the statements are conditional on monitoring and expertise playing some role in these markets. Whether these factors are sufficiently important to explain asset pricing puzzles quantitatively in those markets is more speculative and left to future research.

The rest of the paper proceeds as follows. Section 1 reviews the literature. Section 2 presents the simplest possible asset pricing model without risk aversion, to get at the main results of the paper. Section 3 endogenizes intermediaries' financial capacity, by allowing them to lever up. Section 4 builds a general equilibrium model with an endogenous interest rate and an "occupational choice" for investors, who choose to become borrowers or lenders. Section 5 discusses the applications of the model. Section 6 concludes.

# 1 Literature

Asset pricing based on risk aversion, or the mean-variance portfolio optimization, has a long tradition in theoretical asset pricing (Markowitz (1952), Sharpe (1964), Lintner (1965)). A very large literature in theoretical and empirical asset pricing has been concerned with explaining the data through an appropriate choice of price processes and investors' preferences, so that the risk-return works quantitatively. The successes and limitations of this research agenda are discussed in Cochrane (2011) and Cochrane (2017). In this paper, I choose to abstract from these forces altogether, because the main message of the paper is qualitative and conceptual.

The theory put forward in this paper is very closely related to intermediary-based asset pricing. This theory notes that the marginal investor, for whom a risk-return trade-off is operative, is typically a financial institution, whose stochastic discount factor should therefore be determining asset prices. Therefore, the risk-return trade-off should be applied to financial institutions' preferences, rather than to households' preferences. This literature includes He and Krishnamurthy (2013), Brunnermeier and Sannikov

(2014). Adrian et al. (2014) and He et al. (2017) have shown that this theory is much more successful at pricing the cross-section of assets. One open question in this literature is to understand why intermediaries don't hedge with consumers, as Krishnamurthy (2003) and Di Tella (2017) have pointed out.

The paper also contributes to the large body of literature in corporate finance, corporate governance, and financial frictions, surveyed in Tirole (2006). This literature is very rich, with Bernanke and Gertler (1989), Shleifer and Vishny (1992), Kiyotaki and Moore (1997), Holmstrom and Tirole (1997), Holmström and Tirole (1998), and Bernanke et al. (1999). The modeling in this paper is very similar to Moll (2014), in particular with regards the assumption of linear technologies and heterogeneous entrepreneurs. To the best of my knowledge, the asset pricing implications of these theories when lenders are heterogeneous in efficiency, or second-best value uses are different, have not been studied in these models. This addition of heterogeneity to financial frictions models is the main contribution of the paper. The Liquidity Asset Pricing Model (LAPM) by Holmström and Tirole (2001), has also looked at the implications of models of liquidity for asset pricing, but from a liquidity premium perspective.

Finally, the paper borrows methods from hedonic pricing following Rosen (1974), as the skin in the game parameter $\lambda$ plays the role of a priced characteristic. To the best of my knowledge, this paper is the first which looks at the implication of models of compensating differentials, and two-sided sorting models, for asset pricing in rational models. The model is close to Geerolf (2015), who has investigated a similar sorting model with behavioral agents entertaining different beliefs in the tradition of Lucas (1978). The present model does not have a fixed asset supply, and the asset pricing implications of the model are unambiguous, unlike in Geerolf (2015), where realized returns may be too high or too low, depending on the bias in the belief distribution.

# 2 A stylized model of asset pricing without risk aversion

## 2.1 Basic setup

The model is in discrete time, and has two periods: $t = 0, 1$. There are two types of agents: borrowers ("entrepreneurs"), and lenders ("bankers"). To stay as close as possible to the literature on financial frictions, I assume that the two sides of the market are predetermined: there are lenders on the one hand, and borrowers on the other hand. Appendix D looks at a similar model with occupational choice: the results that follow do not rely on ex-ante heterogeneous behavior.

**Lenders.** A mass of potential lenders ("loan officers"), born with wealth 1, allow entrepreneurs to pledge a fraction $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ per invested unit in the project, with $\bar{\lambda} < \underline{z}$.

This pledgeability, or skin in the game, parameter determines what fraction of the unit investment may be pledged to outside investors. There is a density $f(.)$ of these technologies, so that the number of lenders with pledgeability parameters in $[\lambda, \lambda + d\lambda]$ is $f(\lambda)d\lambda$.

A key assumption that distinguishes this paper from the literature on financial frictions is that pledgeability parameters $\lambda$ are heterogenous across lenders. Several underlying frictions can give rise to such an agency wedge $\lambda$. These frictions are the subject of an important literature surveyed in Tirole (2006), and we shall take this pledgeability parameter as given here.[2] The most straightforward interpretation is perhaps one where lenders have heterogeneous recovery values, given by $\lambda$. If $\lambda$ is the value of the asset for the lender, then the borrower will always be able to renegotiate the payment up to the lenders' own valuation of the asset. Anticipating this, a lender will only lend up to that valuation, that is, if a lender is able to extract a value $\lambda$ when left with the asset, he will lend only $\lambda$.

Other interpretations of underlying financial frictions equally give rise to lender heterogeneity. For example, in a limited commitment model, a borrower may be able to "run away with" (or "steal") an amount $1 - \lambda$ per unit of asset. In this interpretation, $1 - \lambda$ could depend on the borrower that the lender faces. A higher $\lambda$ banker ("a better banker") might be able to hold on to assets more efficiently, so that the lender may run away with a smaller fraction of the asset.

**Borrowers.** There is also a continuum of borrowers ("entrepreneurs"), who are endowed with a productivity level $z \in [\underline{z}, \bar{z}]$, born with wealth 1, which is without loss of generality.[3] This productivity level $z$ determines the number of units of output generated in period 1 when one unit of wealth is invested in period 0. The number of entrepreneurs in $[z, z + dz]$ is given by $g(z)dz$. There is an outside investment technology with a return equal to $1 + R$, which is exogenously given – this assumption will be relaxed in Appendix 4. I assume that entrepreneurs have a more productive use than the outside investment technology so that $\underline{z} > 1 + R$. However, the heterogeneity on the side of borrowers is not needed for the main results, particularly for the asset pricing implications of the model – see appendix B. That is, we could assume homogenous $z$, with $\underline{z} = \bar{z}$, as in Appendix B.

[2] These include limited commitment, private benefits of shirking, second-best use of the asset, incentives to counter moral hazard, or recovery values in bankruptcy. This way of modeling the financing friction is very standard in the literature; I refer the reader to Tirole (2006) for multiple justifications of this agency friction, as well as for the microfoundations of more complex contracting structures - such as block-holding, etc.

[3] The reason for why this is without loss of generality should be clear later. All technologies and constraints will be linear in wealth, and there will be a continuous distribution of entrepreneurs defined by a density function $f(z)$. This density function could therefore represent both how many entrepreneurs there are at productivity $z$ and also how much wealth they have.

**Markets.** Markets are competitive. Lenders and borrowers take prices as given. Both lenders' and borrowers' types are observable. In this stylized model, lenders are endowed with some wealth but may not use outside investment. This assumption is made to keep as close as possible to textbook models of financing frictions (Tirole (2006)). This assumption is relaxed in section 3.

## 2.2 Assumptions

Before proceeding to the equilibrium of this model, two assumptions are next made: one is substantial, and the other one is technical.

**Non-zero heterogeneity among lenders.** First, I assume that there is a non-zero heterogeneity among lenders: that is, the heterogeneity across lenders is not zero. This assumption is one distinguishes this paper from the rest of the literature on financial frictions, and so one cannot dispense with it. In terms of the primitives of the model, this implies that the density function $g(.)$ is not the density of a Dirac point distribution at $\bar{\lambda} = \underline{\lambda}$.

**Assumption 1.** *The distribution of pledgeability parameters $\lambda$ is non degenerate.*

As already mentioned, a similar assumption is not needed for borrowers, who need not be heterogenous.

**Abundant lending capacity.** While the previous assumption is a substantial assumption, this assumption is only a technical assumption. Anticipating the results in the paper, I define the interest rate on the loan $r(\lambda)$ as the amount $\lambda$ promised in period 1 divided by the loan amount $q(\lambda)$ that borrowers get in period 0, when they promise $\lambda$. This interest rate is given by:

$$r(\lambda) = \frac{\lambda}{q(\lambda)}.$$

Depending on the relative aggregate initial endowments of lenders and borrowers, lenders with the lowest pledgeability parameter will be investing in the storage technology or lending to entrepreneurs. If they are lending to entrepreneurs, then entrepreneurs with the lowest productivity will not be able to lever up. Expected returns would be even higher in this case because entrepreneurs would be competing for scarce lenders' resources. In the following, I therefore make the following assumption, which implies that there is no excess return on the lowest leverage ratio loan.

**Assumption 2.** *Lenders have sufficient funds so that the outside technology is used in equilibrium, which pins down the interest rate on the lowest leverage ratio loan:*

$$r(\underline{\lambda}) = 1 + R.$$

Unlike Assumption 1, Assumption 2 is not crucial for the main results. Again, the reader is invited to refer to Appendix B for the case of homogeneous borrowers.

### 2.3 Equilibrium

Proposition 1 characterizes the equilibrium of the economy defined in section 2.3. The implications, and its relationship to textbook asset pricing, is discussed in section 2.4.

**Proposition 1.** *(Equilibrium) In equilibrium, there exists a threshold $\lambda_m$ for the skin in the game parameter, such that lenders with $\lambda \geq \lambda_m$ lend to entrepreneurs, lenders with $\lambda \leq \lambda_m$ invest in the outside technology. There is positive sorting between lenders and entrepreneurs, such that lenders with the highest pledgeability parameter $\lambda$ lend to the most productive entrepreneurs. Denoting by $q(\lambda)$ the price of promising $\lambda$ in period 1, and by $z(\lambda)$ defined over $[\lambda_m, \bar{\lambda}]$ the increasing function matching lenders to entrepreneurs, then $\lambda_m$, $q(\lambda)$ and $z(\lambda)$ satisfy the following set of two first-order ordinary differential equations, and three algebraic equations:*

$$(z(\lambda) - \lambda)\, q'(\lambda) = 1 - q(\lambda) \tag{1}$$

$$q(\lambda) g\left(z(\lambda)\right) z'(\lambda) = \left(1 - q(\lambda)\right) f(\lambda) \tag{2}$$

$$(a) \quad \lambda_m = (1+R) q(\lambda_m), \qquad (b) \quad z(\bar{\lambda}) = \bar{z}, \qquad (c) \quad z(\lambda_m) = \underline{z} \tag{3}$$

*Proof.* Proposition 1 is proved in several steps. Assumption 2 implies equation (3a):

$$\lambda_m = (1+R) q(\lambda_m)$$

The fact that a competitive equilibrium leads lenders with a relatively high skin in the game parameter to lend to entrepreneurs is intuitive, as they allow to scale entrepreneurs' investment upwards and leads to a higher productivity. Note also that given positive sorting, the lenders with the highest (lowest) pledgeability parameter lend to the entrepreneurs with the highest (lowest) productivity, which implies equations (3b) and (3c):

$$z(\bar{\lambda}) = \bar{z} \quad \text{and} \quad z(\lambda_m) = \underline{z}.$$

**Borrowers' problem.** At time $t = 1$, the entrepreneur can produce $z$ per unit of asset invested in period 0. Since he promised to repay $\lambda$ in period 1, his profit is $z - \lambda$ per unit of asset invested. The balance sheet for each unit of investment is represented on

Figure 9.

The size of the investment is also endogenous. The entrepreneur wants to invest the maximum possible amount in his project - that is, each unit of wealth that he borrows will be put to productive use in the project, none will be invested in storage. When borrowing from a lender with a pledgeability parameter $\lambda$, and denoting by $q(\lambda)$ the price of such funds, he will therefore need to finance $1 - q(\lambda)$ of the project through his own funds. His profit in period 1 will therefore be equal to $(z - \lambda)/(1 - q(\lambda))$. The balance sheet for each unit of wealth is represented on Figure 10.

The problem of the borrower will therefore be to choose the lender he is borrowing from optimally to maximize his return on equity:

$$\max_{\lambda} \quad \frac{z - \lambda}{1 - q(\lambda)}.$$

Entrepreneurs' return can be decomposed as follows:

$$\begin{aligned}
\frac{z - \lambda}{1 - q(\lambda)} &= \left(1 + \frac{q(\lambda)}{1 - q(\lambda)}\right) z - \frac{\lambda}{1 - q(\lambda)} \\
&= z + \left[z - \frac{\lambda}{q(\lambda)}\right] \frac{q(\lambda)}{1 - q(\lambda)} \\
\frac{z - \lambda}{1 - q(\lambda)} &= z * 1 + [z - r(\lambda)] \underbrace{\frac{q(\lambda)}{1 - q(\lambda)}}_{\text{Borrowed Funds}}.
\end{aligned}$$

If the problem is interior, then this implies:

$$-(1 - q(\lambda)) + q'(\lambda)(z - \lambda) = 0 \quad \Rightarrow \quad \frac{z - \lambda}{1 - q(\lambda)} = \frac{1}{q'(\lambda)}.$$

There is a straightforward economic intuition behind this optimality condition. In equilibrium, the benefits of borrowing from a lender with a higher pledgeability parameter are equal to the costs. For simplicity, it is simplest to think of this trade-off per unit of investment, since the entrepreneurs' problem is linear. It is also simplest to think in term of income at time $t = 1$. Borrowing from a marginally more productive lender ($\lambda + d\lambda$ rather than $\lambda$, with $d\lambda > 0$) allows the entrepreneur to raise $dq(\lambda)$ additional units of funds. In equilibrium, the return from these additional funds are equal to the entrepreneurs' expected return, and therefore raising $dq(\lambda)$ more in period 0 allows to get $dq(\lambda) * (z - \lambda)/(1 - q(\lambda))$ more in period 1. The cost comes from having to repay more in period 1, $d\lambda$ more:

$$\underbrace{dq(\lambda)}_{\substack{\text{Extra Funds}\\ t=0}} \quad \underbrace{\frac{z-\lambda}{1-q(\lambda)}}_{\substack{\text{Return on wealth}\\ \text{t=0 to t=1}}} = \underbrace{d\lambda.}_{\substack{\text{Extra Repayment}\\ \text{t=1}}} \quad \Rightarrow \quad \frac{z-\lambda}{1-q(\lambda)} = \frac{1}{q'(\lambda)}.$$

This proves equation (1).

**Positive sorting.** From the return on wealth obtained by an entrepreneur with productivity $z$, borrowing from lenders with pledgeability parameter $\lambda$, which is also equal to the quantity of output produced $Q(\lambda, z)$, given by:

$$Q(\lambda, z) = \frac{z}{1-q(\lambda)}.$$

There is a positive cross-partial of this amount with respect to the productivity of the entrepreneur and that of the lender as:

$$\frac{\partial^2 Q(\lambda, z)}{\partial \lambda \partial z} = \frac{q'(\lambda)}{(1-q(\lambda))^2} > 0$$

where $q'(\lambda) > 0$ comes from entrepreneurs' optimization problem. No entrepreneur would ever be willing to borrow from a more productive lender, who requires a higher amount of repayment ($\lambda$), if that lender did not as well provide more funds ex-ante. There is positive sorting in equilibrium. Denote by $z(\lambda)$ the matching between lenders and entrepreneurs. That is, for each productivity level of lenders $\lambda$, the matching function associates the productivity of the entrepreneur it lends to (all entrepreneurs of such type are identical, so it is sufficient to describe the equilibrium). In an interval $[\lambda, \lambda + d\lambda]$ of such lenders, the number of units of investment that they allow entrepreneurs to buy is given by $g(\lambda)d\lambda/q(\lambda)$, by definition of the density function. The number of units of investment that entrepreneurs can buy in an interval $[z, z+dz]$ is in turn given by $f(z)dz/\left(1-q(\lambda)\right)$. Thus, finally the resource constraint gives a first order differential equation as follows:

$$\frac{g(\lambda)d\lambda}{q(\lambda)} = \frac{f(z)dz}{1-q(\lambda)} \quad \Rightarrow \quad z'(\lambda) = \frac{1-q(\lambda)}{q(\lambda)} \frac{g(\lambda)}{f\left(z(\lambda)\right)}.$$

This proves equation (2).

□

A first-order ordinary differential equation and an algebraic equation is an initial value problem. Thus, proposition 1 contains two initial value problems and one algebraic equation, which jointly determine $\lambda_m$, $q(\lambda)$ and $z(\lambda)$ .

### 2.4 Implications for asset pricing

Proposition 1 is the main result of the paper. This proposition derives an equilibrium with heterogenous lenders and borrowers, in which heterogenous interest rates are market clearing prices between borrowers and lenders. In this theory, asset pricing is neither determined by different expected cash flows, nor by risk aversion, but corresponds to the scarcity of various monitoring technologies. One can state a number of corollaries.

**No arbitrage relations.** A first corollary is that no arbitrage relationships seem violated in this model: indeed, the model is deterministic, so that there is only one state of the world. Yet different securities, which lead to similar cash flows in this only state of the world, have different expected returns. From the point of view of traditional finance theory, this would be considered a violation of no-arbitrage.

**Corollary 1.** *In equilibrium, there is a violation of no arbitrage, according to the usual finance definition: similar cash flows have different prices in the same state of the world. There is no unique price for Arrow-Debreu securities.*

Corollary 1 is important because the assumption because no arbitrage is an important component of finance theory. For example, this assumption is crucial in demonstrating the existence of a stochastic discount factor.[4]

At the same time, this is not a feasible arbitrage in the context of the model, although it would be defined as an arbitrage using the usual definition. A high $\lambda$ lender has his wealth already fully committed in high $\lambda$ contracts, and the extreme financial constraint he is facing does not allow him to take a larger position. In contrast, a low $\lambda$ agent could in principle sell his low $\lambda$ contracts and buy high $\lambda$ contracts, which would allow him to earn a higher expected return. However given that his technology is in fact of a lower $\lambda$ type, he would not be able to get the full value from such a contract, so that his return would not be $r(\lambda)$.

What is key for apparent violation of arbitrage is that the econometrician might not take into account that different securities have different implied skin in the game, and that therefore they cannot be bought by any investor. According to this theory, a rationale for an apparent arbitrage opportunities comes from unobserved heterogeneity in $\lambda$.

**Comparative statics: lenders' and borrowers' aggregate wealth.** A second implication of proposition 1 is that expected returns are negatively related to lenders'

[4]For example, a definition of no arbitrage is given in Cochrane (2009): "A payoff space X and pricing function $p(x)$ leave no arbitrage opportunities if every payoff $x$ that is always nonnegative, $x \geq 0$ (almost surely), and positive, $x > 0$, with some positive probability, has positive price, $p(x) > 0$." No arbitrage is defined as a constraint on the payoff space such that all positive payoffs must have a positive price. In particular, no arbitrage implies that there is a law of one price for Arrow-Debreu securities.

aggregate wealth.

Rhese comparative static exercises only rely on supply and demand for funds with different pledgeability parameters (different $\lambda$). It does not rely on intermediaries becoming more risk averse as they lose wealth – which relies on some assumptions.

At the same time, the theory has several asset pricing implications in common with traditional intermediary-based asset pricing, and can therefore relate to much of the evidence supportive of this theory. Quite coincidentally, it remains true that assets with a higher equilibrium covariance with intermediary leverage have higher expected returns on average. However, the mechanism does not rely on intermediaries' risk aversion, but simply on supply and demand for financial expertise.

**Comparative statics: lenders' individual wealth.** A third implication of proposition 1. This is consistent with Collin-Dufresne et al. (2001), who document that monthly credit spread changes on corporate bonds "are primarily driven by local supply/demand shocks that are independent of both credit-risk factors and standard proxies for liquidity." These comparative statics are more standard, even though they are quite different from intermediary-based asset pricing.

**Corollary 2.** *(Comparative statics with respect to other lenders' wealth) The model also predicts that in fact, everyone's wealth, both lenders' and entrepreneurs", have impacts on expected return. In particular, this means that shocks that happen to other parts of the financial sector should have an impact on assets that unaffected banks hold.*

## 2.5 Illustration with a uniform density

In the case of a uniform density function, the equations in Proposition 1 have a closed-form solution. I solve for this case first, since the message of the paper is mostly conceptual. I then study the comparative statics of this model, illustrating the results using this particular density function.

Section 2.4 reflects further on these comparative static exercises, which are very different from those derived in textbook asset pricing. What matters for asset pricing in this model is the dispersion of productivities, the dispersion of monitoring technologies, the relative wealth of borrowers and lenders, etc.

**Lemma 1.** *(Uniform density) Assume a uniform density of lenders over* $[\underline{\lambda}, \bar{\lambda}]$ *given by* $g(\lambda) = \Lambda$*, and a uniform density of entrepreneurs over* $[\underline{z}, \bar{z}]$ *given by* $f(z) = Z$*. Define:*

$$a \equiv \frac{g(\lambda)}{f(z)} = \frac{\Lambda}{Z}$$

*Then the interest rate function and the matching function are given as a function* *of $C_1$ and $C_2$ by:*

$$r(\lambda) = \frac{\lambda}{C_2\left[(1+a)\lambda - C_1\right]^{\frac{1}{1+a}}}, \qquad z(\lambda) = \lambda + \frac{1}{C_2}\left[(1+a)\lambda - C_1\right]^{\frac{a}{1+a}} - \left[(1+a)\lambda - C_1\right],$$

*where $C_1$,$C_2$ and $\lambda_m$ are solutions of a system of three equations:*

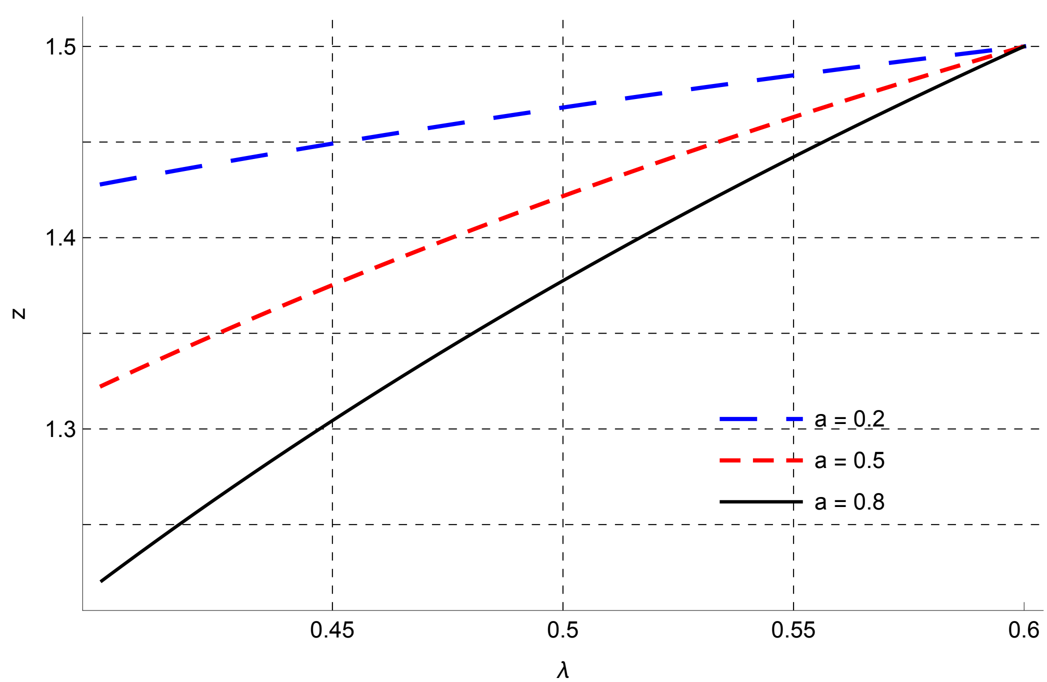
$$\begin{aligned}
\lambda_m =& (1+R)C_2\left[(1+a)\lambda_m - C_1\right]^{\frac{1}{1+a}} \\
\underline{z} =& \lambda_m + \frac{1}{C_2}\left[(1+a)\lambda_m - C_1\right]^{\frac{a}{1+a}} - \left[(1+a)\lambda_m - C_1\right] \\
\bar{z} =& \bar{\lambda} + \frac{1}{C_2}\left[(1+a)\bar{\lambda} - C_1\right]^{\frac{a}{1+a}} - \left[(1+a)\bar{\lambda} - C_1\right]
\end{aligned}$$

*Proof.* The derivation of these expressions uses elementary algebra, starting from two first-order ordinary differential equations (1) and (2), which are solved up to two constants $C_1$ and $C_2$. Solving these two equations in closed form relies on a change of variable involving the natural logarithm of $q(\lambda) = \lambda/r(\lambda)$, denoted by $u(\lambda) = \log(q(\lambda))$. $C_1$, $C_2$ and $\lambda_m$ are then a solution to equations (3a), (3b) and (3c). Appendix A.1 provides all the details. □

**Comparative statics.** Figure 1 plots the matching function $z(\lambda)$, with the following parameters values:

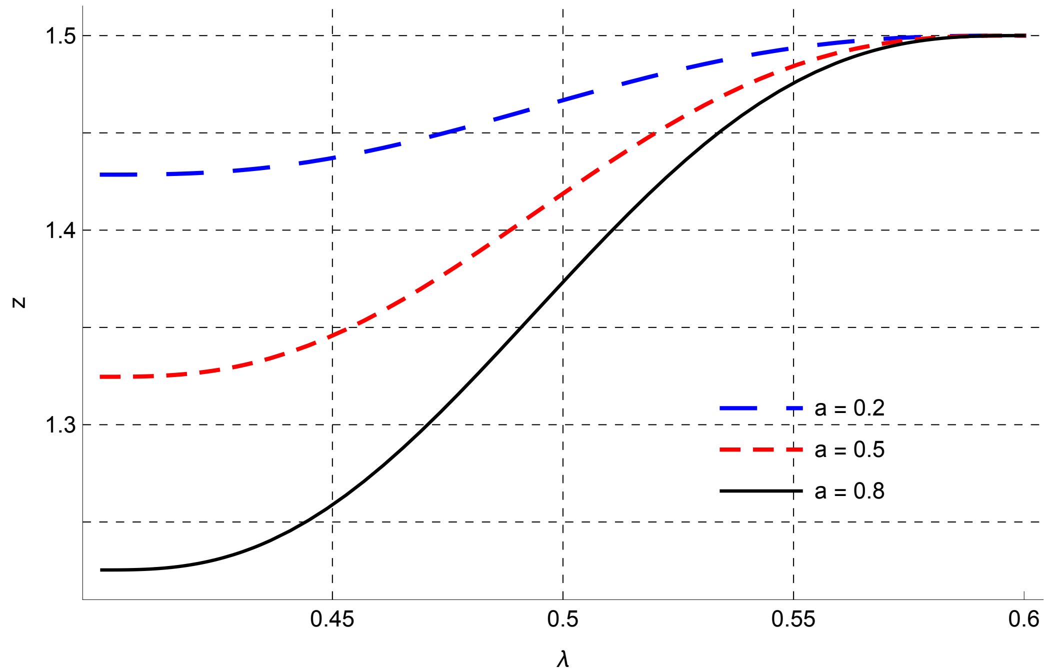
$$\underline{\lambda} = 0.4, \quad \bar{\lambda} = 0.6, \quad \underline{z} = 1.2, \quad \bar{z} = 1.5.$$

Figure 1: Matching Function, $f$ = Uniform, $g$ = Uniform



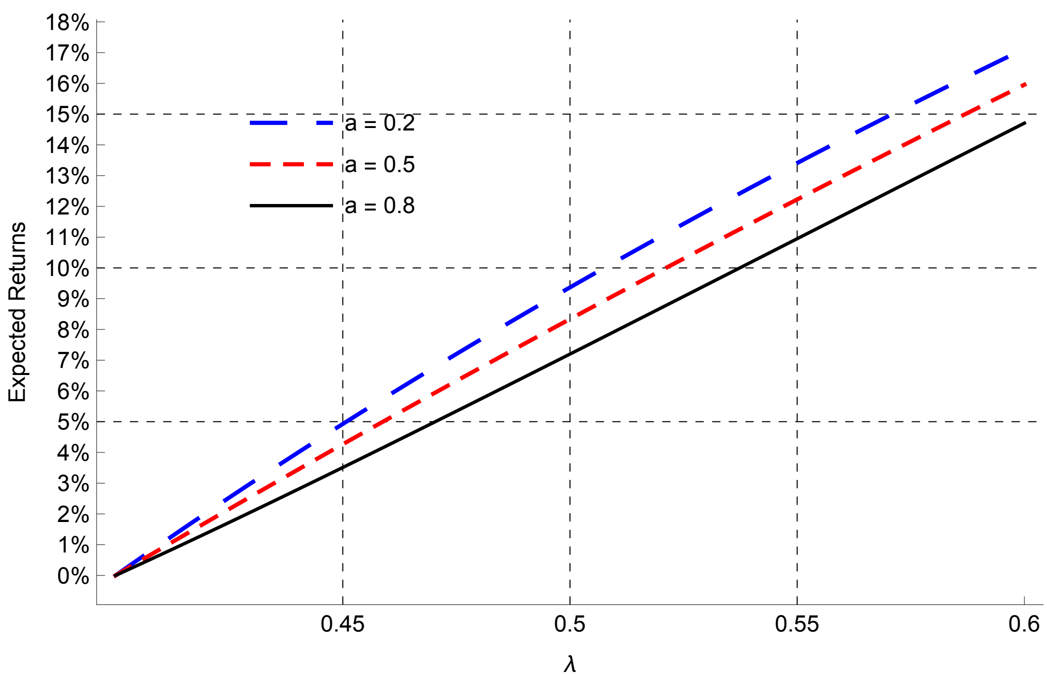

**Note:** The matching function from lenders' efficiency $\lambda$ to entrepreneurs' productivity $z = z(\lambda)$ is plotted for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 2: Matching Function, $f$ = Beta(2,2), $g$ = Uniform



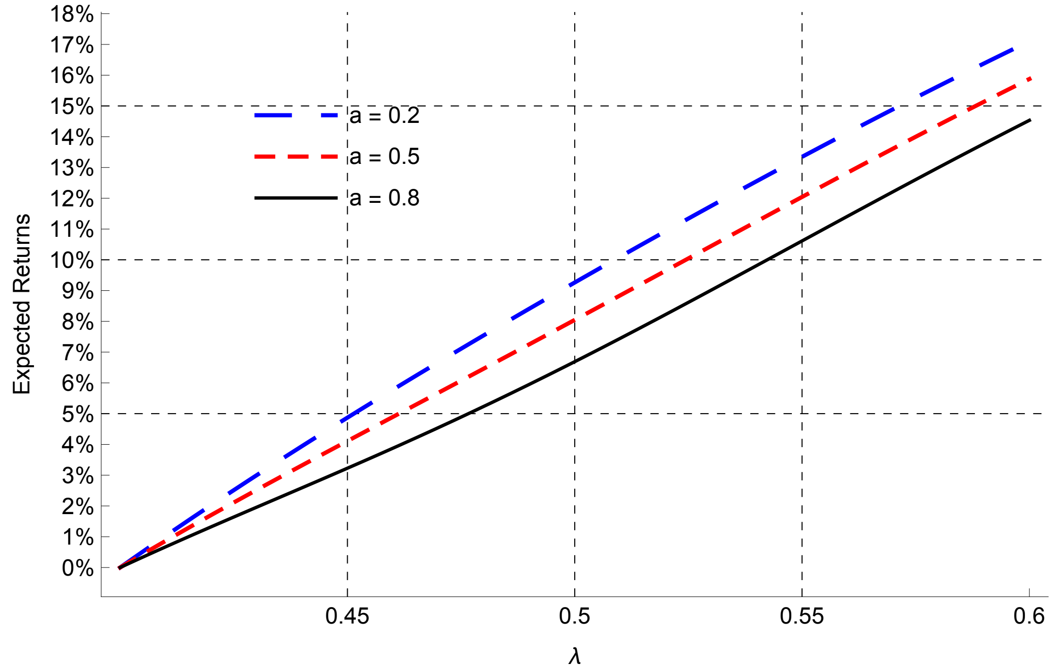

**Note:** The matching function from lenders' efficiency $\lambda$ to entrepreneurs' productivity $z = z(\lambda)$ is plotted for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 3: Expected Returns, $f$ = Uniform, $g$ = Uniform



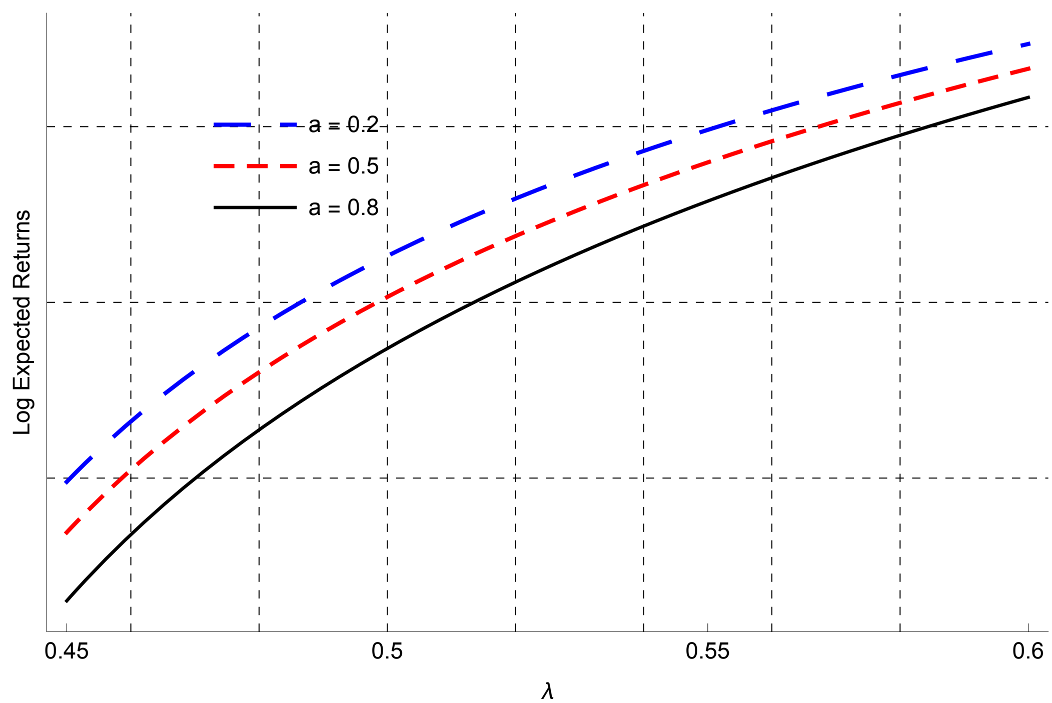

**Note:** Interest rate spreads $r(\lambda) - 1$ are plotted as a function of loans' face value for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 4: Expected Returns, $f$ = Beta(2,2), $g$ = Uniform



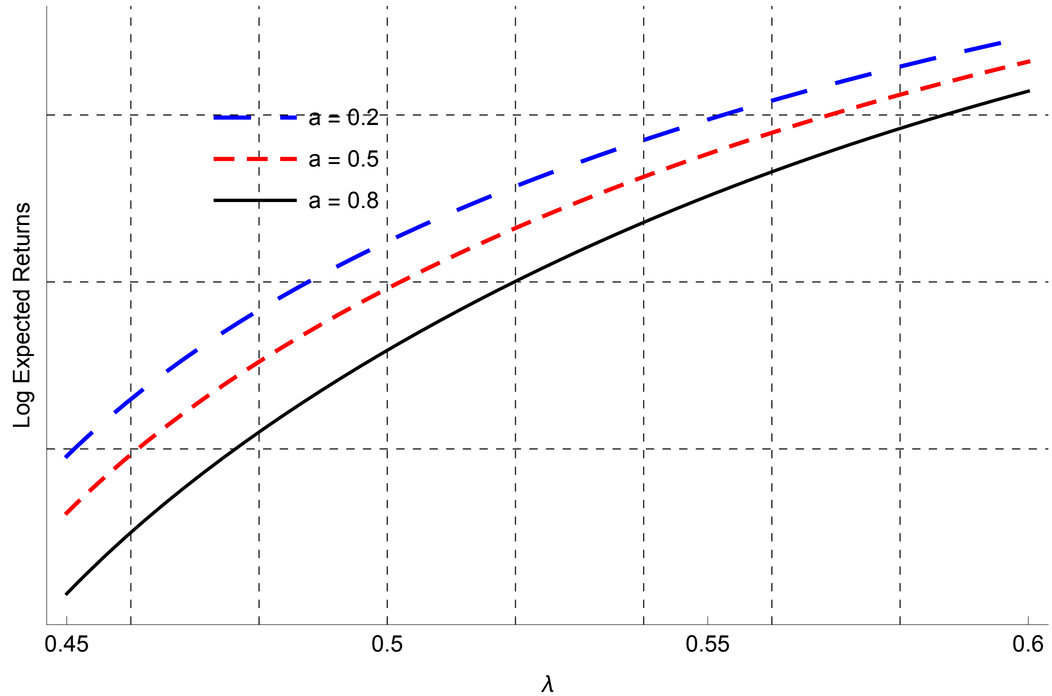

**Note:** Interest rate spreads $r(\lambda) - 1$ are plotted as a function of loans' face value for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 5: Log Expected Returns, $f$ = Uniform, $g$ = Uniform



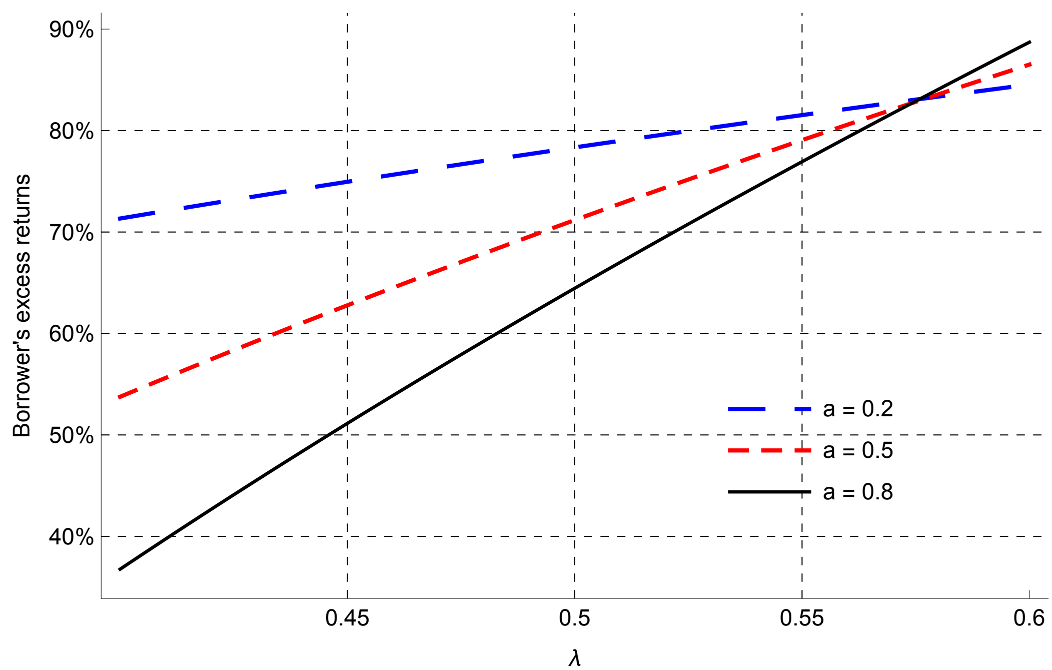

**Note:** Log Interest rate spreads $\log(r(\lambda) - 1)$ are plotted for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 6: Log Expected Returns, $f$ = Beta(2,2), $g$ = Uniform



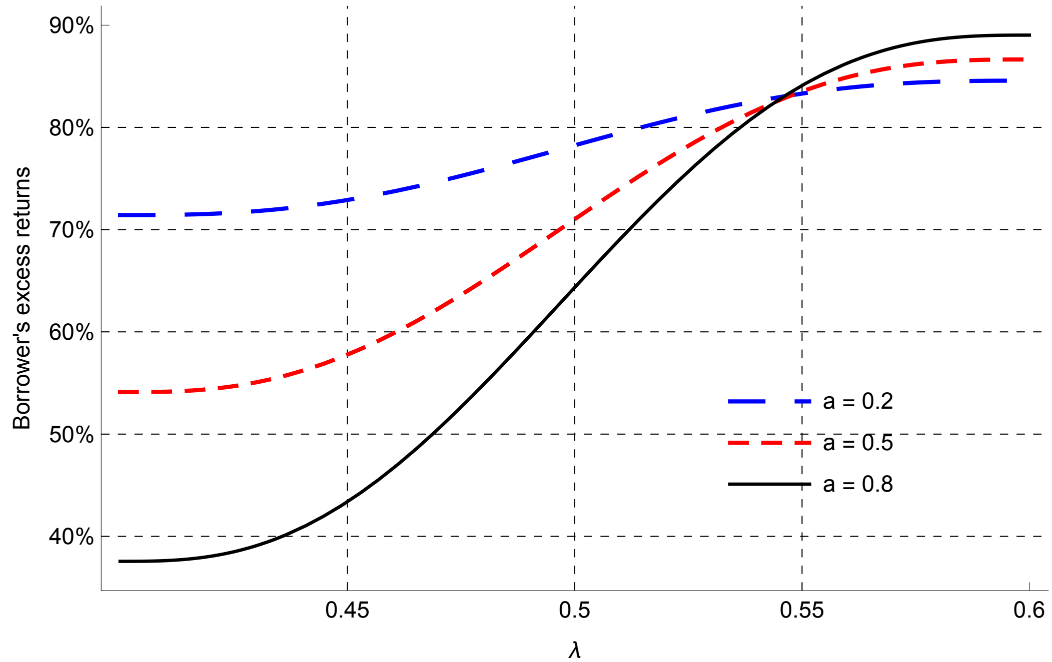

**Note:** Log Interest rate spreads $\log(r(\lambda) - 1)$ are plotted for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 7: Borrowers' excess returns, $f$ = Uniform, $g$ = Uniform




**Note:** Borrowers' excess returns $(z(\lambda)-\lambda)/(1-q(\lambda))-1$ are plotted for the following parameter values: $\underline{\lambda} = 0.4$, $\bar{\lambda} = 0.6$, $\underline{z} = 1.2$, $\bar{z} = 1.5$, and different values of $a$.

Figure 8: Borrowers' excess returns, $f$ = Beta(2,2), $g$ = Uniform




**Note:** Borrowers' excess returns $(z(\lambda)-\lambda)/(1-q(\lambda))-1$ are plotted for the following parameter values: $\underline{\lambda}=0.4$, $\bar{\lambda}=0.6$, $\underline{z}=1.2$, $\bar{z}=1.5$, and different values of $a$.

### 2.6 Scarce labor or time

In the setup presented in section 2.1, lenders can only lend up to their initial endowment. The next section provides one foundation for such a constraint: lenders may be subject to the same financial frictions as entrepreneurs. However, another type of constraint may perhaps simply arise from the fact that financial expertise is not just a rent, but also requires the financier to supply some labor.

Although some elasticity of substitution could be assumed between money and labor, one can assume a Leontief technology for illustration, which measures the number of units of financial services that are provided. This Leontief technology would then reflect how much financial expertise, as well as how much labor, must be provided in order to produce financial services:

$$y(\lambda, l_\lambda) = \min\{\lambda, l_\lambda\}$$

With such a technology, provided that available capital does not bind, then capital has a zero marginal return if more labor is not provided. Thus, one reason why intermediaries with a monitoring technology equal to $\lambda$ may be constrained, is simply that this expertise is embodied in individuals who have only limited time or labor (think of loan officers for example, who have a lot of experience).

The model thus may provide a new interpretation for fire sales. In Shleifer and Vishny (1992), fire sales rely on decreasing marginal values of second-best uses as more assets are sold on the market. Another line of research has investigated the role of search frictions in reallocating assets to natural buyers in times of high selling pressure (Weill (2007)). This "slow moving capital" (Duffie (2010)) may explain why in many markets, such as catastrophe risk insurance, the dynamic of premia are mainly driven by shocks to the capital levels of insurers. The present model provides a new way to think about why capital does not immediately flow to apparent trading opportunities.

# 3 Endogenous intermediary leverage

In 2, I have exogenously imposed that lenders could not lever up. In this section, I relax this assumption and endogenize lenders' leverage.

**Proposition 2.** *(Equilibrium) In equilibrium, there exist two cutoffs $\lambda_{1m}$, $\lambda_{2m}$ for the pledgeability parameter, two matching functions $z(\lambda)$ defined over $[\lambda_{1m}, \bar{\lambda}]$, and $\lambda(\lambda')$ defined over $[\lambda_{2m}, \lambda_{1m}]$ and two bond pricing functions defined over $[\lambda_{1m}, \bar{\lambda}]$ and $[\lambda_{2m}, \lambda_{1m}]$ such that:*

- *Potential lenders are split into three intervals:*
  - *Potential lenders with $\lambda \in [\underline{\lambda}, \lambda_{2m}]$ invest in the outside technology with $R = 1$.*
  - *Potential lenders with $\lambda \in [\lambda_{2m}, \lambda_{1m}]$ lend to lenders to entrepreneurs.*
  - *Potential lenders with $\lambda \in [\lambda_{1m}, \bar{\lambda}]$ lend directly to entrepreneurs.*
- *There is positive sorting between lenders and entrepreneurs, and lenders to lenders and lenders, such that lenders with the highest pledgeability parameter $\lambda$ lend to the most productive entrepreneurs, and lenders to lenders with the highest pledgeability parameter lend to the lenders with the highest pledgeability parameter. These correspondence from $\lambda$ to $z$ and $\lambda'$ to $\lambda$ are denoted $z(\lambda)$ and $\lambda'(\lambda)$. Positive sorting implies: $z'(\lambda) > 0$ and $(d\lambda/d\lambda')(\lambda') > 0$.*
- *The matching functions $z(\lambda)$, $\lambda(\lambda')$ , and the bond pricing functions $q_1(\lambda)$ and $q_2(\lambda')$ are such that:*

$$(z(\lambda) - \lambda)\, q_1'(\lambda) = 1 - q_1(\lambda) \tag{4}$$

$$q_1(\lambda) g\,(z(\lambda))\, z'(\lambda) = (1 - q_1(\lambda))\, f(\lambda) \tag{5}$$

$$(\lambda - \lambda') q_2'(\lambda') = q_1(\lambda) - q_2(\lambda') \tag{6}$$

$$q_2(\lambda') f\,(\lambda)\, \frac{d\lambda}{d\lambda'}(\lambda') = \big(q_1(\lambda) - q_2(\lambda')\big)\, f(\lambda') \tag{7}$$

$$q_2(\lambda_{2m}) = \lambda_{2m}, \quad z(\bar{\lambda}) = \bar{z}, \quad z(\lambda_{1m}) = \underline{z},$$
$$\frac{\lambda_{1m}}{q_2(\lambda_{1m})} = \frac{\bar{\lambda} - \lambda_{1m}}{q_1(\bar{\lambda}) - q_2(\lambda_{1m})}, \quad \lambda(\lambda_{2m}) = \lambda_{1m}, \quad \lambda(\lambda_{1m}) = \bar{\lambda}.$$

Proposition 2 is proved in several steps. Assumption 2 implies directly:

$$q_2(\lambda_{2m}) = \lambda_{2m}$$

Given positive sorting, the lenders with the highest (lowest) pledgeability parameter lend to the entrepreneurs with the highest (lowest) productivity:

$$z(\bar{\lambda}) = \bar{z} \quad \text{and} \quad z(\lambda_{1m}) = \underline{z}.$$

There is also positive sorting between lenders to lenders and lenders, so that:

$$\lambda(\lambda_{2m}) = \lambda_{1m} \quad \text{and} \quad \lambda(\lambda_{1m}) = \bar{\lambda}.$$

Indifference for agents with pledgeability parameter $\lambda_{1m}$:

$$\frac{\lambda_{1m}}{q_2(\lambda_{1m})} = \frac{\bar{\lambda} - \lambda_{1m}}{q_1(\bar{\lambda}) - q_2(\lambda_{1m})}.$$

**Borrowers' problem.** Entrepreneurs' problem is given by:

$$\max_{\lambda} \quad \frac{z - \lambda}{1 - q_1(\lambda)}.$$

If the problem is interior, then this implies:

$$-(1 - q_1(\lambda)) + q_1'(\lambda)(z - \lambda) = 0 \quad \Rightarrow \quad \frac{z - \lambda}{1 - q_1(\lambda)} = \frac{1}{q_1'(\lambda)}.$$

**Lenders' problem.** Lenders' problem is given by:

$$\max_{\lambda'} \quad \frac{\lambda - \lambda'}{q_1(\lambda) - q_2(\lambda')}$$

If the problem is interior, then this implies:

$$-\left(q_1(\lambda) - q_2(\lambda')\right) + q_2'(\lambda')(\lambda - \lambda') = 0 \quad \Rightarrow \quad \frac{\lambda - \lambda'}{q_1(\lambda) - q_2(\lambda')} = \frac{1}{q_2'(\lambda')}.$$

**Positive sorting.** There is positive sorting between entrepreneurs and lenders, but also between lenders lending directly to entrepreneurs, and lenders lending to lenders. The resource constraint gives a first order differential equation as follows:

$$\frac{f(\lambda)d\lambda}{q_1(\lambda)} = \frac{f(z)dz}{1 - q_1(\lambda)} \quad \Rightarrow \quad z'(\lambda) = \frac{1 - q_1(\lambda)}{q_1(\lambda)} \frac{f(\lambda)}{f(z(\lambda))}.$$

A second resource constraint similarly equates:

$$\frac{f(\lambda')d\lambda'}{q_2(\lambda')} = \frac{f(\lambda)d\lambda}{q_1(\lambda)-q_2(\lambda')} \quad \Rightarrow \quad \frac{d\lambda}{d\lambda'}(\lambda') = \frac{q_1(\lambda)-q_2(\lambda')}{q_2(\lambda')}\frac{f(\lambda')}{f(\lambda)}.$$

# 4 General Equilibrium: occupational choice and endogenous safe interest rate

## 4.1 Setup

The setup is similar to that which is described in Section 2, except that borrowers and lenders decide on borrowing and lending endogenously, depending on their respective productivities. The model has two periods: $t = 0, 1$. There is a continuum of agents, born with different productivities $z \in [\underline{z}, \bar{z}]$. There is an outside investment technology with a return $1 + R$, such that $(1 + R) \in ]\underline{z}, \bar{z}[$. The number of agents with productivity z is given by $f(z)dz$. When they become entrepreneurs, agents operate with productivity $z$. When they lend instead, $z$ stands in for the pledgeability parameter. A very natural interpretation is that conditional on default, the lender would be constrained to operate the asset himself. Therefore, how much the borrower can credibly pledge to the lender is what the lender's own productivity is at operating the asset (Hart and Moore (1994)). Finally, I assume that the distribution of productivities has full support on $[\underline{z}, \bar{z}]$. Proposition D characterizes the equilibrium with occupational choice.

The setup is similar to that described in Section D, except that the outside storage technology is public debt, whose amount sold is exogenously given by $B$ (one can think for example that the government needs to raise a given amount of funds). I assume that the amount of public debt sold is lower than the total amount of total available saving:

$$B \leq \int_{\underline{z}}^{\bar{z}} f(z)dz.$$

As previously, the amount of saving does not depend on the interest rate.

## 4.2 Equilibrium

**Proposition.** *(General equilibrium) There exist two thresholds $\lambda_m$ and $z_m$ for productivity, such that agents with $z \geq z_m$ become borrowers, agents with $\lambda_m \leq z \leq z_m$ become lenders and agents with $\lambda \leq \lambda_m$ invest in public debt. There is positive sorting between lenders and entrepreneurs, such that lenders with the highest pledgeability parameter $\lambda$ lend to the most productive entrepreneurs. Denoting by $1 + R$ the endogenous interest rate on public debt, $q(\lambda)$ the price of promising $\lambda$ in period 1, and by $z(\lambda)$ defined over $[\lambda_m, z_m]$ the increasing function matching lenders to entrepreneurs, then $\lambda_m$, $z_m$, $R$,*

*$q(\lambda)$ and $z(\lambda)$ satisfy the following set of equations:*

$$\begin{aligned}
&(z(\lambda) - \lambda)\, q'(\lambda) = 1 - q(\lambda) \\
&q(\lambda) f\left(z(\lambda)\right) z'(\lambda) = \left(1 - q(\lambda)\right) f(\lambda) \\
&\textit{(a)} \quad \lambda_m = (1+R) q(\lambda_m), \quad \textit{(b)} \quad z(z_m) = \bar{z}, \\
&\textit{(c)} \quad z(\lambda_m) = z_m, \quad \textit{(d)} \quad z_m = \frac{z_m - \lambda_m}{1 - q(\lambda_m)} q(z_m), \quad \textit{(e)} \quad \int_{\underline{z}}^{\lambda_m} f(z) dz = B.
\end{aligned} \tag{8}$$

*Proof.* The ordinary differential equations, as well as equations (8a), (8b), (8c) and (8d) can be derived as for Proposition D. Equation (8b) results from market clearing for saving: agents with $z \in [\underline{z}, \lambda_m]$ invest in public debt. □

### 4.3 Comparative statics

**Uniform density.** Again, the case of a uniform density function allows to get some intuition for the model. Denote $f(z) = Z$. Then, the matching function $z(\lambda)$ and the pricing function $q(\lambda)$ are given by:

$$\begin{aligned}
q(\lambda) &= C_2 \sqrt{(1+a)\lambda - C_1}, \\
z(\lambda) &= \lambda + \frac{1}{C_2} \sqrt{(1+a)\lambda - C_1} - \left[(1+a)\lambda - C_1\right].
\end{aligned}$$

Finally, $C_1$, $C_2$, $\lambda_m$, $z_m$, and $R$ solve equations (8a), (8b), (8c), (8d) and (8e) with $q(.)$ and $z(.)$ given by these two expressions.

### 4.4 Asset pricing implications

**Comparative statics with respect to the dispersion in returns.** The comparative statics exercise with respect to $B$ in the model correspond to the exercise in Krishnamurthy and Vissing-Jorgensen (2012), which is replicated on Figure 17 and on Figure 18. They show that the spread between AAA corporate bonds and Treasuries is a decreasing function of the amount of public debt. They interpret it as a money demand curve: the less public debt there is, the more valued it becomes and therefore the more expensive it is (public debt is more expensive when its yield is further below the return on AAA corporate bonds). This model provides a different interpretation. In this model, a higher level of public debt implies less dispersion in expected returns in general, because entrepreneurs' returns are not as different when fewer of them produce (only the most productive are able to produce).

**Comparative statics with respect to the safe interest rate $R$.** When $B$ increases, the safe interest rate $R$ increases, which is standard. In typical models, it arises from a declining marginal product of capital. Here in contrast, there is a linear technology but there are financial constraints and heterogeneous entrepreneurs, so that the marginal project declines in productivity as more capital is being invested.

# 5 Applications

I next provide some candidate applications of the above asset pricing model to different segments of the financial markets. Section 5.1 looks at closely held investments, such as private equity and venture capital. Section 5.3 relates the findings in the model to those that have been previously documented for mutual funds. The last two sections are more speculative, and more work is needed to investigate whether asset pricing might indeed be determined by the scarcity of expertise in all of these markets. Section 5.4 interprets the evidence in Adrian et al. (2014) about public equity markets, which has been related to intermediary-based asset pricing, in the light of this model. Finally, section 5.5 similarly relates the behavior of corporate bonds to intermediary-based asset pricing, as in He et al. (2017).

## 5.1 Private equity and venture capital

A first piece of evidence, which is supportive of the main mechanisms highlighted in the paper comes from the private equity market, and particularly from Venture Capital (VC) funds. This is probably where investors' monitoring is most visible. According to Hall and Woodward (2010), "general partners (...) arrange financing and supervise the startup company by holding board seats."

**Fact 1.** *(Heterogeneity and persistence in returns across VC investors) Substantial heterogeneity in returns across venture capital funds exist, and persist over time (Kaplan and Schoar (2005)).*

This observation by Kaplan and Schoar (2005) is what the stylized model presented in section 2 predicts. Higher $\lambda$ investors tend to finance more productive entrepreneurs, who sell claims to their future payoff $z$ given by $\lambda$ at a lower price $q(\lambda)$ and therefore earn a higher expected return $r(\lambda)$. To the extent that $\lambda$ is a persistent characteristic of VC investors, then these returns will also be persistent. Moreover, Kaplan and Schoar (2005) provide an explanation which is in line with the intuition developed in this paper. They assert that "underlying heterogeneity in the skill and quality of General Partners could lead to heterogeneity in performance and to more persistence if new entrants cannot compete effectively with existing funds." They also conjecture that "a startup would be willing to accept to these terms if some investors provided superior management, advisory, or reputational inputs." Similarly, Bottazzi et al. (2008) report that "venture capital firms with partners that have prior business experience are more active recruiting managers and directors, helping with fundraising, and interacting more frequently with their portfolio companies." Finally, Sørensen (2007) shows that experienced Venture Capitalists invest in better companies. These comments correspond to the main economic forces at play in the model of section 2.

**Fact 2.** *(Better Venture Capitalists have higher expected returns) Better venture capitalists get better deal terms (lower valuations) when negotiating with startups. (Hsu (2004))*

Hsu (2004) show that better Venture Capitalists get better deal terms (lower valuations) when negotiating with startups. According to Hsu (2004), high-reputation Venture Capitalists acquire start-up equity at a 10-14% discount. Once again, these empirical facts are very consistent with the model presented in section 2. Indeed, higher $\lambda$ investors do acquire equity at a discount compared to the initial face value, given by $q(\lambda) = \lambda/r(\lambda)$ where the discount is given by the equilibrium expected return $r(\lambda)$.

## 5.2 Syndicated loans

The model allows to rationalize practitioners' observation that credit spreads mostly depend on the basic laws of supply and demand in the market, which have no obvious counterpart in traditional asset pricing theory (where prices should only depend on the stochastic discount factor of the representative investor. For example, in Standard and Poor's a Guide to the Loan market: "In pricing loans to institutional investors, it's a matter of the spread of the loan relative to credit quality and market-based factors. This second category can be divided into liquidity and market technicals (i.e., supply/demand)/ Market technicals, or supply relative to demand, is a matter of simple economics. If there are a lot of dollars chasing little product, then, naturally, issuers will be able to command lower spread. If, however, the opposite is true, then spread will need to increase for loans to clear the market."

## 5.3 Mutual funds

A second potential application of the model concerns mutual fund returns. The closest model to this market structure is presented in Appendix E. The assumption there would be that mutual funds are the experts, and that investors in mutual funds are uninformed investors. For example, mutual funds may not just pick stocks but can also vote in shareholder's meetings and are actively monitoring the companies they invest in – which can be more indirect, such as voting with activist hedge funds. Then the gross of fees performance of mutual funds could also be coming from this return to financial expertise. Their investments would bring them more returns, as an equilibrium reward to their higher $\lambda$. Berk and Green (2004) describe the fact that mutual funds are highly compensated "one of the central mysteries facing financial economics."

**Fact 3.** *(High compensation of mutual funds) Mutual funds are highly rewarded, despite the apparent fierce competition between them. (Berk and Green (2004))*

By contrast, the model presented in this paper would suggest that these rewards are driven by fair market prices for scarce expert services.

Moreover, Carhart (1997) note that some mutual funds benefit from momentum strategies, not because they actively pick last year's winners, but because they happen to have a high exposure in last year winning stocks .

**Fact 4.** *Funds that earn higher one-year returns do so not because fund managers successfully follow momentum strategies, but because some mutual funds just happen by chance to hold relatively larger positions in last year's winning stocks. (Carhart (1997))*

If one believes that expertise is somewhat persistent, then some expert mutual funds would hold relatively larger positions in assets that have higher expected returns, just because they are the source of these higher expected returns. They would hold on to that portfolio, because monitoring is also persistent.

## 5.4 Public equity markets

A third application concerns the public equity markets. I start from Adrian et al. (2014), who document the ability of intermediary asset pricing model to price equity portfolios (book to market, size, momentum) and ask what assumptions in the model may allow to microfound this correlation.

**Fact 5.** *(Intermediary leverage factor and Fama and French (1992) portfolios) Higher covariance of Fama and French (1992) equity portfolios (book-to-market and size) with broker dealer leverage is correlated with higher expected returns (Adrian et al. (2014)).*

The question is whether stocks of small firms and high book to market firms are more likely to be held by experts. This is possible, as it could come from higher financial constraints or higher importance of monitoring when firms are "distressed". Then the model would be able to explain why these firms have higher returns, and why their stock prices comove with intermediaries' leverage. This is a testable proposition. Note that the view that only intermediaries can buy the Fama and French (1992) portfolios is implicit in Adrian et al. (2014)'s approach.

**Fact 6.** *(intermediary leverage factor and Jegadeesh and Titman (1993) momentum) Higher covariance of Jegadeesh and Titman (1993) momentum with broker dealer leverage is correlated with higher expected returns (Adrian et al. (2014)).*

The hypothesis that expected returns reflect the price of expertise is also consistent with observing a momentum effect in stocks. Indeed, if monitoring is persistent, then stocks that are monitored will persistently earn higher expected returns. For the same reason, these stocks' expected returns will be more correlated with broker dealer leverage, and so it will look like intermediary leverage is a priced factor. Once again, Adrian et al. (2014)'s evidence on the link between momentum and the intermediary leverage factor can be rationalized through the model.

Note also that the Fama and French (2015) 5-factor model includes investment and profitability factors. The explanatory power of the intermediary leverage factor has however not been established for these portfolios. However, one can hypothesize that stocks with higher profitability have higher expected returns, because of the positive sorting result in Proposition 1. Similarly, the observation that stocks of firms with more conservative investment behavior have higher expected returns can be rationalized in the context of the present model. These firms are likely more financially constrained – and thus, their financing is subject to the type of mechanisms which have been put forward in this paper. These questions are left to future research.

### 5.5 Corporate bonds and sovereign debt

Finally, motivated by the evidence in favor of intermediary based asset pricing shown by He et al. (2017), I now discuss the valuation of corporate and sovereign bonds.

**Fact 7.** *(Intermediary leverage factor and corporate bonds) Higher covariance of corporate bond prices with intermediary leverage is correlated with higher expected returns (He et al. (2017)).*

Again, the following statements are somewhat speculative and conditional on monitoring and expertise playing some role in these markets. In the case of corporate bond investors, the corresponding special type of expertise could consist in approaching the distressed firms and giving them advice, which would be akin to monitoring services. For defaulted corporate bonds, it might involve having vulture investors buying corporate bonds at discounted prices (see Lewis (2016)). Similarly, for sovereign bonds, one could think that vulture hedge funds would have different abilities to sue defaulting governments in various jurisdictions, and thereby have different expected recovery values on the same bonds.

This interpretation of corporate bonds could potentially help explain the puzzling observation from Collin-Dufresne et al. (2001), who document that monthly credit spread changes on corporate bonds "are primarily driven by local supply/demand shocks that are independent of both credit-risk factors and standard proxies for liquidity." Again, the model in section 2 gives a role for supply/demand shocks to drive expected returns. But once again, more research is required to shed light on these issues.

# 6 Conclusion

This paper has developed a theory of asset pricing without risk aversion, in which asset prices reflect the relative prices of scarce monitoring technologies. In this model, cash is valued differently depending on the financial constraints that are attached to it. Expected returns on one unit of wealth are heterogeneous depending on monitoring technologies, and thus, who holds the asset. In such a model, a financial economist who does not understand that there exists different monitoring technologies across agents, might wrongly conclude on violations of arbitrage. Equivalent promises of cash at the same time and in similar states of the world have different expected returns, only because these assets are held by investors with different monitoring technologies.

The purpose of this paper was mostly qualitative and conceptual. However, the heterogeneity in expected returns has been shown to have the same order of magnitude as the dispersion of productivities. From earlier work on misallocation and productivity heterogeneity, these order of magnitudes may be substantial, perhaps in the order of 30%. I leave a proper calibration to future work.

Section 5 has provided direct evidence that these asset pricing effects are present in the private equity and venture capital markets. In this industry, it is well-known that better venture capital funds do get better deal terms, which means that indeed, expertise is priced in the form of higher expected returns. I have also discussed the cases of mutual funds, public equity markets, as well as corporate and sovereign debt, where the evidence is much less direct, but where the relevance of the theory is plausible. More research needs to be done in order to determine in which of these excess returns indeed correspond to the returns to financial expertise.

## References


**Adrian, Tobias, Erkko Etula, and Tyler Muir**, "Financial Intermediaries and the Cross-Section of Asset Returns," *The Journal of Finance*, 2014, *69* (6), 2557–2596. 2, 5, 6, 23, 25, 26, 41

**Berk, Jonathan B. and Richard C. Green**, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, December 2004, *112* (6), 1269–1295. 24, 25

**Bernanke, Ben and Mark Gertler**, "Agency Costs, Net Worth, and Business Fluctuations," *The American Economic Review*, 1989, *79* (1), 14–31. 6

**Bernanke, Ben S., Mark Gertler, and Simon Gilchrist**, "Chapter 21 The financial accelerator in a quantitative business cycle framework," in "Handbook of Macroeconomics," Vol. 1, Elsevier, January 1999, pp. 1341–1393. 6

**Bottazzi, Laura, Marco Da Rin, and Thomas Hellmann**, "Who are the active investors?," *Journal of Financial Economics*, September 2008, *89* (3), 488–512. 23

**Brunnermeier, Markus K. and Yuliy Sannikov**, "A Macroeconomic Model with a Financial Sector," *American Economic Review*, February 2014, *104* (2), 379–421. 2, 5

**Carhart, Mark M.**, "On Persistence in Mutual Fund Performance," *The Journal of Finance*, March 1997, *52* (1), 57–82. 25

**Cochrane, John H.**, *Asset Pricing (Revised Edition)*, Princeton University Press, April 2009. Google-Books-ID: 20pmeMaKNwsC. 12

_ , "Presidential Address: Discount Rates," *The Journal of Finance*, August 2011, *66* (4), 1047–1108. 2, 5

_ , "Macro-Finance," *Review of Finance*, May 2017, *21* (3), 945–985. 2, 5

**Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin**, "The Determinants of Credit Spread Changes," *The Journal of Finance*, December 2001, *56* (6), 2177–2207. 13, 26, 44

**Duffie, Darrell**, "Presidential Address: Asset Price Dynamics with Slow-Moving Capital," *The Journal of Finance*, August 2010, *65* (4), 1237–1267. 2, 19

**Fama, Eugene F. and Kenneth R. French**, "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, June 1992, *47* (2), 427–465. 2, 25

**Fama, Eugene F and Kenneth R French**, "A five-factor asset pricing model," *Journal of Financial Economics*, 2015, *116* (1), 1–22. 26

**Geerolf, François**, "Leverage and Disagreement," *Working Paper*, 2015, pp. 1–61. 6

**Hall, Robert E. and Susan E. Woodward**, "The Burden of the Nondiversifiable Risk of Entrepreneurship," *American Economic Review*, June 2010, *100* (3), 1163–1194. 23

**Hart, Oliver and John Moore**, "A Theory of Debt Based on the Inalienability of Human Capital," *The Quarterly Journal of Economics*, 1994, *109* (4), 841–879. 21, 35

**He, Zhiguo and Arvind Krishnamurthy**, "Intermediary Asset Pricing," *American Economic Review*, April 2013, *103* (2), 732–770. 2, 5

— **, Bryan Kelly, and Asaf Manela**, "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, October 2017, *126* (1), 1–35. 2, 5, 6, 23, 26

**Holmstrom, Bengt and Jean Tirole**, "Financial Intermediation, Loanable Funds, and The Real Sector," *The Quarterly Journal of Economics*, August 1997, *112* (3), 663–691. 6

**Holmström, Bengt and Jean Tirole**, "Private and Public Supply of Liquidity," *Journal of Political Economy*, February 1998, *106* (1), 1–40. 6

— **and** — , "LAPM: A Liquidity-Based Asset Pricing Model," *The Journal of Finance*, October 2001, *56* (5), 1837–1867. 6

**Hsu, David H.**, "What Do Entrepreneurs Pay for Venture Capital Affiliation?," *The Journal of Finance*, August 2004, *59* (4), 1805–1844. 24

**Jegadeesh, Narasimhan and Sheridan Titman**, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *The Journal of Finance*, March 1993, *48* (1), 65–91. 25

**Kaplan, Steven N. and Antoinette Schoar**, "Private Equity Performance: Returns, Persistence, and Capital Flows," *The Journal of Finance*, August 2005, *60* (4), 1791–1823. 4, 23

**Kiyotaki, Nobuhiro and John Moore**, "Credit Cycles," *Journal of Political Economy*, April 1997, *105* (2), 211–248. 6

**Krishnamurthy, Arvind**, "Collateral constraints and the amplification mechanism," *Journal of Economic Theory*, August 2003, *111* (2), 277–292. 6

— **and Annette Vissing-Jorgensen**, "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, April 2012, *120* (2), 233–267. 22, 43

— **and Tyler Muir**, "How Credit Cycles across a Financial Crisis," Working Paper 23850, National Bureau of Economic Research September 2017. 2

**Lewis, Ryan**, "Corporate Debt Markets and Recovery Rates with Vulture Investors," SSRN Scholarly Paper ID 2539585, Social Science Research Network, Rochester, NY January 2016. 26

**Lintner, John**, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *The Review of Economics and Statistics*, 1965, *47* (1), 13–37. 5

**Lucas, Robert E.**, "Asset Prices in an Exchange Economy," *Econometrica*, 1978, *46* (6), 1429–1445. 6

**Markowitz, Harry**, "Portfolio Selection," *The Journal of Finance*, March 1952, *7* (1), 77–91. 5

**Moll, Benjamin**, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, October 2014, *104* (10), 3186–3221. 6

**Muir, Tyler**, "Financial Crises and Risk Premia," *The Quarterly Journal of Economics*, May 2017, *132* (2), 765–809. 2, 42

**Rosen, Sherwin**, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 1974, *82* (1), 34–55. 6

**Sharpe, William F.**, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, 1964, *19* (3), 425–442. 5

**Shleifer, Andrei and Robert W. Vishny**, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *The Journal of Finance*, September 1992, *47* (4), 1343–1366. 6, 18

**Sørensen, Morten**, "How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital," *The Journal of Finance*, December 2007, *62* (6), 2725–2762. 24

**Tella, Sebastian Di**, "Uncertainty Shocks and Balance Sheet Recessions," *Journal of Political Economy*, August 2017. 6

**Tirole, Jean**, *The Theory of Corporate Finance*, Princeton University Press, 2006. Google-Books-ID: p683L1gaVm4C. 6, 7, 8

**Weill, Pierre-Olivier**, "Leaning Against the Wind," *The Review of Economic Studies*, October 2007, *74* (4), 1329–1354. 19

# A Proofs

## A.1 Proof of Lemma 1

*Proof.* Assuming:

$$\frac{f(\lambda)}{f(z(\lambda))} = \frac{\Lambda}{Z} = a$$

Equation (2) of Proposition 1 implies:

$$z'(\lambda) = \frac{1-q(\lambda)}{q(\lambda)}a$$

Equation (1) implies:

$$q'(\lambda) = \frac{1-q(\lambda)}{z(\lambda)-\lambda} \quad \Rightarrow \quad z(\lambda)-\lambda = \frac{1-q(\lambda)}{q'(\lambda)} \quad \Rightarrow \quad z'(\lambda) = 1 + \frac{-q'(\lambda)^2 - q''(\lambda)(1-q(\lambda))}{q'(\lambda)^2}$$

$$\Rightarrow \quad \frac{1-q(\lambda)}{q(\lambda)}a = -\frac{q''(\lambda)(1-q(\lambda))}{q'(\lambda)^2} \quad \Rightarrow \quad q''(\lambda)q(\lambda) + aq'(\lambda)^2 = 0.$$

thus, using $q$ as a shorthand for function $q(\lambda)$:

$$q''q + aq'^2 = 0 \quad \Rightarrow \quad \frac{q''}{q} + a\frac{q'^2}{q^2} = 0.$$

Using the change of variable $u \equiv \log(q)$:

$$u' = \frac{q'}{q}, \quad u'' = \frac{q''q - q'^2}{q^2} = \frac{q''}{q} - u'^2$$

Therefore:

$$\frac{q''}{q} = -au'^2 \quad \Rightarrow \quad u'' = -(a+1)u'^2$$

then, the previous differential equation simplifies into:

$$u'' + (a+1)u'^2 = 0 \quad \Rightarrow \quad -\frac{u''}{u'^2} = a+1 \quad \Rightarrow \quad \frac{1}{u'} = (a+1)\lambda - C_1$$

$$\Rightarrow \quad u' = \frac{1}{(a+1)\lambda - C_1} \quad \Rightarrow \quad u(\lambda) = \frac{1}{1+a}\log\left[(a+1)\lambda - C_1\right] + \log C_2$$

$$\Rightarrow \quad q(\lambda) = \exp\left[u(\lambda)\right] = C_2\left[(1+a)\lambda - C_1\right]^{\frac{1}{1+a}}$$

Expected returns are thus given in closed form by:

$$r(\lambda) = \frac{\lambda}{q(\lambda)} = \frac{\lambda}{C_2\left[(1+a)\lambda - C_1\right]^{\frac{1}{1+a}}}.$$

The $q'(\lambda)$ function in the uniform case is:

$$q'(\lambda) = C_2\left[(1+a)\lambda - C_1\right]^{\frac{1}{1+a}-1}$$

Note that the matching function $z(.)$ can be written only as a function of $q(.)$:

$$\begin{aligned} z(\lambda) &= \lambda + \frac{1-q(\lambda)}{q'(\lambda)}. \\ &= \lambda + \frac{1 - C_2\left[(1+a)\lambda - C_1\right]^{\frac{1}{1+a}}}{C_2\left[(1+a)\lambda - C_1\right]^{\frac{1}{1+a}-1}} \\ z(\lambda) &= \lambda + \frac{1}{C_2}\left[(1+a)\lambda - C_1\right]^{\frac{a}{1+a}} - \left[(1+a)\lambda - C_1\right]. \end{aligned}$$

Finally, $C_1$,$C_2$ and $\lambda_m$ are given by substituting in equations (3a) , (3b) and (3c):

$$\begin{aligned}
q(\lambda_m) = \lambda_m \quad &\Rightarrow \quad C_2\left[(1+a)\lambda_m - C_1\right]^{\frac{1}{1+a}} = \lambda_m \\
z(\lambda_m) = \underline{z} \quad &\Rightarrow \quad \lambda_m + \frac{1}{C_2}\left[(1+a)\lambda_m - C_1\right]^{\frac{a}{1+a}} - \left[(1+a)\lambda_m - C_1\right] = \underline{z} \\
z(\bar{\lambda}) = \bar{z} \quad &\Rightarrow \quad \bar{\lambda} + \frac{1}{C_2}\left[(1+a)\bar{\lambda} - C_1\right]^{\frac{a}{1+a}} - \left[(1+a)\bar{\lambda} - C_1\right] = \bar{z}.
\end{aligned}$$

This proves Lemma 1. $\square$

## A.2 Example: $a = 1$

Assuming $\Lambda = Z$, then $a = 1$. From equation (3a) and substituting $a = 1$:

$$C_2\sqrt{2\lambda_m - C_1} = \lambda_m \quad \Rightarrow \quad C_1 = 2\lambda_m - \left(\frac{\lambda_m}{C_2}\right)^2.$$

From equation (3b):

$$\begin{aligned}
\lambda_m + \frac{1}{C_2}\sqrt{2\lambda_m - C_1} - (2\lambda_m - C_1) = \underline{z} \quad &\Rightarrow \quad \lambda_m + \frac{1}{C_2}\frac{\lambda_m}{C_2} - \left(\frac{\lambda_m}{C_2}\right)^2 = \underline{z} \\
&\Rightarrow \quad C_2^2 = \frac{\lambda_m - \lambda_m^2}{\underline{z} - \lambda_m} \quad \Rightarrow \quad C_2 = \sqrt{\frac{\lambda_m - \lambda_m^2}{\underline{z} - \lambda_m}}.
\end{aligned}$$

Substituting in the previous equation:

$$C_1 = 2\lambda_m - \frac{\underline{z} - \lambda_m}{1 - \lambda_m}\lambda_m \quad \Rightarrow \quad C_1 = \frac{2\lambda_m - \lambda_m^2 - \underline{z}}{1 - \lambda_m}.$$

Therefore:

$$2\bar{\lambda} - C_1 = 2\left(\bar{\lambda} - \lambda_m\right) + \frac{\underline{z} - \lambda_m}{1 - \lambda_m}\lambda_m$$

From equation (3c):

$$\begin{aligned}
&\bar{\lambda} + \frac{1}{C_2}\sqrt{2\bar{\lambda} - C_1} - \left(2\bar{\lambda} - C_1\right) = \bar{z} \quad \Rightarrow \quad 2\bar{\lambda} - C_1 - \frac{1}{C_2}\sqrt{2\bar{\lambda} - C_1} + \bar{z} - \bar{\lambda} = 0 \\
&\Rightarrow \quad 2\left(\bar{\lambda} - \lambda_m\right) + \frac{\underline{z} - \lambda_m}{1 - \lambda_m}\lambda_m + \sqrt{\frac{\underline{z} - \lambda_m}{\lambda_m - \lambda_m^2}}\sqrt{2\left(\bar{\lambda} - \lambda_m\right) + \frac{\underline{z} - \lambda_m}{1 - \lambda_m}\lambda_m} + \bar{z} - \bar{\lambda} = 0.
\end{aligned}$$

$\lambda_m$ has a closed form solution from this last equation. Since $C_1$ and $C_2$ are closed form functions of $\lambda_m$, so are they. Thus $r(\lambda)$ and $z(\lambda)$ also have closed form solutions.

## A.3 Example: $a = 0$

Assuming $Z = +\infty$, then $a = 0$. From equation (3a) and substituting $a = 0$:

$$C_2\left(2\lambda_m - C_1\right) = \lambda_m \quad \Rightarrow \quad C_1 = 2\lambda_m - \frac{\lambda_m}{C_2}.$$

From equation (3b):

$$1 + \frac{1}{C_2} - (2\lambda_m - C_1) = \bar{z} \quad \Rightarrow \quad 1 + \frac{1}{C_2} - \frac{\lambda_m}{C_2} = \bar{z} \quad \Rightarrow \quad C_2 = \frac{1 - \lambda_m}{\bar{z} - 1}.$$

Substituting in the previous equation:

$$C_1 = 2\lambda_m - \frac{\bar{z} - 1}{1 - \lambda_m}\lambda_m \quad \Rightarrow \quad C_1 = \frac{2\lambda_m - \bar{z} + 1}{1 - \lambda_m}\lambda_m.$$

Therefore:

$$2\bar{\lambda} - C_1 = 2\left(\bar{\lambda} - \lambda_m\right) + \frac{\bar{z}-1}{1-\lambda_m}\lambda_m.$$

From equation (3c):

$$\bar{\lambda} + \frac{1}{C_2} - \left(2\bar{\lambda} - C_1\right) = \bar{z} \quad \Rightarrow \quad \bar{\lambda} + \frac{\bar{z}-1}{1-\lambda_m} - \left(2\left(\bar{\lambda} - \lambda_m\right) + \frac{\bar{z}-1}{1-\lambda_m}\lambda_m\right) = \bar{z}$$

$$\Rightarrow \quad \bar{\lambda} + (\bar{z}-1) - 2\left(\bar{\lambda} - \lambda_m\right) = \bar{z} \quad \Rightarrow \quad \lambda_m = \frac{\bar{\lambda}}{2}.$$

$C_1$ and $C_2$ are given as a function of primitives by:

$$C_1 = \frac{\bar{\lambda} - \bar{z} + 1}{2-\bar{\lambda}}\bar{\lambda}, \quad C_2 = \frac{1}{2}\frac{2-\bar{\lambda}}{\bar{z}-1}.$$

# B Homogeneous borrowers

As explained in the introduction of Section 2, the setup described in section 2.1 is not quite the simplest one to get at the main results of the paper. This section presents an even simpler model which does assume any heterogeneity across borrowers. This section illustrates that in fact, the results on expected returns do not rely on sorting.

The setup is similar to that which is described in Section 2. The model has two periods: $t = 0, 1$. There is a continuum of entrepreneurs ("borrowers"), who are endowed with the same productivity level, and all born with wealth 1. The measure of these entrepreneurs is denoted by $\mathcal{Z}$. The corresponding economic object in the setup of section 2 is:

$$\mathcal{Z} \equiv \int_{\underline{z}}^{\bar{z}} f(z)dz.$$

There is an outside investment technology with a return equal to $1 + R$. It is assumed that entrepreneurs have a more productive use than the outside investment technology so that $\underline{z} > 1 + R$. A mass of potential lenders allow entrepreneurs to pledge a fraction $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ per invested unit of the project, with $\bar{\lambda} < 1$. The mass of lenders with skin in the game parameters in $[\lambda, \lambda + d\lambda]$ is given by $f(\lambda)d\lambda$. Assumption 2 holds: lenders have sufficient funds so that the outside technology is used in equilibrium.

**Proposition.** *In equilibrium, lenders with $\lambda \geq \lambda_m$ lend to entrepreneurs, and lenders with $\lambda \leq \lambda_m$ invest in the storage technology. The threshold $\lambda_m$ and pricing function $q(\lambda)$ are given for $\lambda \in [\lambda_m, \bar{\lambda}]$ by the following set of equations:*

$$-(1 - q(\lambda)) + q'(\lambda)(z - \lambda) = 0,$$

$$\mathcal{Z} = \int_{\lambda_m}^{\bar{\lambda}} \frac{1 - q(\lambda)}{q(\lambda)} f(\lambda)d\lambda, \quad \textit{and} \quad (1 + R)q(\lambda_m) = \lambda_m.$$

**Borrowers' problem.** Assuming homogeneous borrowers with productivity $z$, the problem of choosing a lender is given by:

$$\max_{\lambda} \quad \frac{z - \lambda}{1 - q(\lambda)}.$$

This equation implies an indifference condition across $\lambda$: the expected return of a borrower $(z-\lambda)/\left(1 - q(\lambda)\right)$ is constant regardless of his choice of $\lambda$, so that in equilibrium he is indifferent. This indifference condition applies both across $\lambda$ for the same entrepreneur, as well as across entrepreneurs, since they all have the same productivity. Similarly, this leads to a first-order differential equation:

$$-(1 - q(\lambda)) + q'(\lambda)(z - \lambda) = 0 \quad \Rightarrow \quad \frac{z - \lambda}{1 - q(\lambda)} = \frac{1}{q'(\lambda)}.$$

Under Assumption 2:

$$r(\lambda_m) = 1 + R \quad \Rightarrow \quad (1+R)q(\lambda_m) = \lambda_m.$$

Finally, aggregate resources of entrepreneurs are given by $\mathcal{Z}$. When they borrow from a lender who allows them to have a skin in the game $\lambda$, they must provide an amount of funds equal to $1 - q(\lambda)$ for each asset that they buy. How many assets they buy from lenders in interval $[\lambda, \lambda + d\lambda]$ is given by $f(\lambda)d\lambda/q(\lambda)$, the wealth of lenders in that interval divided by the unit price of a loan. Therefore:

$$\mathcal{Z} = \int_{\lambda_m}^{\bar{\lambda}} \frac{1 - q(\lambda)}{q(\lambda)} f(\lambda) d\lambda.$$

This set of equations allows to characterize the equilibrium. Indeed, the above ordinary differential equation can be solved for. Using $q$ as a short notation for $q(\lambda)$ , the above equation can be written as follows:

$$(z - \lambda)q' + q = 1 \quad \Rightarrow \quad \frac{(q-1)'}{q-1} = -\frac{1}{z - \lambda} \quad \Rightarrow \quad (\log(q-1))' = (\log(z - \lambda))'$$
$$\Rightarrow \quad \log\left(\frac{q(\lambda) - 1}{q(\lambda_m) - 1}\right) = \log\left(\frac{z - \lambda}{z - \lambda_m}\right) \quad \Rightarrow \quad q(\lambda) = 1 + \frac{z - \lambda}{z - \lambda_m}(\lambda_m - 1).$$

The return function is given by $r(\lambda) = \lambda/q(\lambda)$. It must therefore be that the expected return of borrowers is constant across $\lambda$. Assuming that entrepreneurs are scarce, then lenders need to be indifferent between the storage technology $1 + R$ and lending to entrepreneurs.

**Uniform density.** The uniform distribution case with $\mathcal{Z} = Z\,(\bar{z} - \underline{z})$ and $f(\lambda) = \Lambda$ can be again calculated in closed form.

[TO BE ADDED]

# C Scarce lenders

Section 2 uses assumption 2 in order to derive an indifference condition for the lowest leverage ratio loan, which is $r(\underline{\lambda}) = 1 + R$. Namely, this amounts to assuming that the scarce side of the market is the side of entrepreneurs. Lenders have more funds than necessary to fund all entrepreneurs fully. Therefore, in equilibrium they compete to lend to entrepreneurs. In this section, I show that this hypothesis is innocuous, and that the results on asset pricing are strengthened if one assumes instead that entrepreneurs are many and that lenders are lacking to fund them. Thus the following assumption is the opposite of assumption 2. It holds if and only if assumption 2 does not hold.

**Assumption.** *Lenders are scarce, so that all entrepreneurs cannot be funded.*

Under that assumption, the least productive entrepreneurs need to be made indifferent between levered investing in their technology and investing in the storage technology. The next proposition is a version of Proposition 1 when the above hypothesis does not hold.

**Proposition.** *(Equilibrium without Assumption 2) In equilibrium, there exists a threshold $z_m$ for borrowers' productivity, such that entrepreneurs with $z \geq z_m$ do levered investing, entrepreneurs with $z \leq z_m$ invest in the outside technology at return $1 + R$. There is positive sorting between lenders and entrepreneurs, such that lenders with the highest pledgeability parameter $\lambda$ lend to the most productive entrepreneurs. Denoting by $q(\lambda)$ the price of promising $\lambda$ in period 1, and by $z(\lambda)$ defined over $[\underline{\lambda}, \bar{\lambda}]$ onto $[z_m, \bar{z}]$ the increasing function that matches lenders to entrepreneurs, then $z_m$, $q(\lambda)$ and $z(\lambda)$ satisfy the following set of two first-order ordinary differential equations, and three algebraic equations:*

$$(z(\lambda) - \lambda)\, q'(\lambda) = 1 - q(\lambda)$$
$$q(\lambda) f\left(z(\lambda)\right) z'(\lambda) = \left(1 - q(\lambda)\right) f(\lambda)$$
$$(a) \quad 1 + R = \frac{z_m - \underline{\lambda}}{1 - q(\underline{\lambda})}, \quad (b) \quad z(\bar{\lambda}) = \bar{z}, \quad (c) \quad z(\underline{\lambda}) = z_m \tag{9}$$

*Proof.* The first-order ordinary differential equations can be derived in the same way as for proposition 1. Sorting at the top still leads to equation (9b). However, sorting at the bottom now involves matching the entrepreneurs with the lowest productivity $z_m$ together with the lenders with the lowest skin in the game parameter $\underline{\lambda}$. Finally, entrepreneurs need to be made indifferent between investing in the outside technology at rate $1+R$ and doing levered investing with their own productivity, which brings:

$$1 + R = \frac{z_m - \underline{\lambda}}{1 - q(\underline{\lambda})}.$$

This proves equation (9c) and completes the proof. □

**Corollary 3.** *Without assumption 2, even the lenders with the lowest pledgeability parameter earn some non-zero spread over the outside technology* $1 + R$*.*

*Proof.* This statement is a consequence of equation (9c), together with the fact that the lowest productivity entrepreneur still is able to do better than the outside technology: $\underline{z} > 1 + R$. Assume that the return of the lenders with the lowest pledgeability parameter was $r(\underline{\lambda}) = 1+R$. Then we would have the following:

$$\frac{z_m - \underline{\lambda}}{1 - \frac{\underline{\lambda}}{1+R}} = 1 + R \quad \Rightarrow \quad z_m = 1 + R.$$

There is a contradiction there, as $z_m > \underline{z} > 1 + R$. Thus, even lenders with $\underline{\lambda}$ earn rents. □

# D Occupational choice

The setup is similar to that described in Section 2, except that borrowers and lenders decide on borrowing and lending endogenously, depending on their respective productivities.
Apart from these differences, the setup is similar to that which is described in Section 2. The model has two periods: $t = 0, 1$. There is a continuum of agents, born with different productivities $z \in [\underline{z}, \bar{z}]$. There is an outside investment technology with a return $1 + R$, such that $(1 + R) \in ]\underline{z}, \bar{z}[$. The number of agents with productivity z is given by $f(z)dz$. When they become entrepreneurs, agents operate with productivity $z$. When they lend instead, $z$ stands in for the pledgeability parameter. A very natural interpretation is that conditional on default, the lender would be constrained to operate the asset himself. Therefore, how much the borrower can credibly pledge to the lender is what the lender's own productivity is at operating the asset (Hart and Moore (1994)). Finally, I assume that the distribution of productivities has full support on $[\underline{z}, \bar{z}]$. Proposition D characterizes the equilibrium with occupational choice.

**Proposition.** *(Equilibrium with occupational choice) There exist two thresholds $\lambda_m$ and $z_m$ for productivity, such that agents with $z \geq z_m$ become borrowers, agents with $\lambda_m \leq z \leq z_m$ become lenders and agents with $\lambda \leq \lambda_m$ invest in the outside technology at rate $1 + R$. There is positive sorting between lenders and entrepreneurs, such that lenders with the highest pledgeability parameter $\lambda$ lend to the most productive entrepreneurs. Denoting by $q(\lambda)$ the price of promising $\lambda$ in period 1, and by $z(\lambda)$ defined over $[\lambda_m, z_m]$ the increasing function matching lenders to entrepreneurs, then $\lambda_m$, $z_m$, $q(\lambda)$ and $z(\lambda)$ satisfy the following set of equations:*

$$\begin{gathered}
(z(\lambda) - \lambda)\, q'(\lambda) = 1 - q(\lambda) \\
q(\lambda) f\left(z(\lambda)\right) z'(\lambda) = \left(1 - q(\lambda)\right) f(\lambda) \\
\textit{(a)} \quad \lambda_m = (1+R) q(\lambda_m), \quad \textit{(b)} \quad z(z_m) = \bar{z}, \\
\textit{(c)} \quad z(\lambda_m) = z_m, \quad \textit{(d)} \quad z_m = \frac{z_m - \lambda_m}{1 - q(\lambda_m)} q(z_m).
\end{gathered} \tag{10}$$

*Proof.* The ordinary differential equations, as well as equation (10a), can be derived as for Proposition 1. The main difference is that now the lenders with the highest skin in the game technology have productivity $z_m$, so that positive sorting at the top implies equation (10b):

$$z(z_m) = \bar{z}.$$

Similarly, positive sorting at the bottom implies equation (10c):

$$z(\lambda_m) = z_m.$$

Finally, because of occupational choice, agents with productivity $z_m$ are indifferent between lending to the most productive entrepreneurs, which gives an expected return equal to $z_m/q(z_m)$, and invest with leverage, borrowing from lender with productivity $\lambda_m$, which brings $(z_m - \lambda_m)/\left(1 - q(\lambda_m)\right)$. This indifference condition leads to equation (10d):

$$z_m = \frac{z_m - \lambda_m}{1 - q(\lambda_m)} q(z_m).$$

□

**Uniform density.** Assuming a uniform density $Z$ on $[\underline{z}, \bar{z}]$, the model can be solved in closed form. The results from A.1 can be used in order to derive the general forms of the pricing function $q(\lambda)$ as well as that of the matching function $z(\lambda)$, with $a = 1$ since the density of entrepreneurs and lenders and collapsed into one. We have the following expressions for $q(\lambda)$ and $z(\lambda)$ :

$$q(\lambda) = C_2 \sqrt{(1+a)\lambda - C_1}$$
$$z(\lambda) = \lambda + \frac{1}{C_2}\sqrt{(1+a)\lambda - C_1} - \left[(1+a)\lambda - C_1\right].$$

$C_1$, $C_2$, $z_m$ and $\lambda_m$ are thus solution to the following set of algebraic equations:

$$\begin{aligned}
\lambda_m =& (1+R) C_2 \sqrt{(1+a)\lambda_m - C_1}, \\
\bar{z} =& z_m + \frac{1}{C_2}\sqrt{(1+a)z_m - C_1} - \left[(1+a)z_m - C_1\right] \\
z_m =& \lambda_m + \frac{1}{C_2}\sqrt{(1+a)\lambda_m - C_1} - \left[(1+a)\lambda - C_1\right] \\
z_m =& \frac{z_m - \lambda_m}{1 - C_2\sqrt{(1+a)\lambda_m - C_1}} C_2 \sqrt{(1+a)z_m - C_1}.
\end{aligned}$$

# E Passive depositors, regulatory constraint

This setup is closer to the setup presented in section 3. However, instead of assuming that lenders to lenders are doing some monitoring, I assume that depositors do not have this technology. Instead, the regulator imposes some exogenous financial constraint on banks, which might for example correspond to its own monitoring technology. In an extension, one could think of having the regulator spend resources to improve that technology, which could allow him to loosen financial regulation and have have the economy produce more.

## F Figures

Figure 9: Section 2: Balance Sheets Per Unit of Asset

| A | L |
|---|---|
| $1$ | $1-q(\lambda)$ |
| | $q(\lambda)$ |

| A | L |
|---|---|
| $z$ | $z-\lambda$ |
| | $\lambda$ |

Figure 10: Section 2: Balance Sheets Per Unit of Wealth

| A | L |
|---|---|
| $\frac{1}{1-q(\lambda)}$ | $1$ |
| | $\frac{q(\lambda)}{1-q(\lambda)}$ |

| A | L |
|---|---|
| $\frac{z}{1-q(\lambda)}$ | $\frac{z-\lambda}{1-q(\lambda)}$ |
| | $\frac{\lambda}{1-q(\lambda)}$ |

Figure 11: Section 3: Balance Sheets Per Unit of Asset

| **A** | **L** |
|---|---|
| $1$ | $1 - q_1(\lambda)$ |
| | $q_1(\lambda)$ |

| **A** | **L** |
|---|---|
| $q_1(\lambda)$ | $q_1(\lambda) - q_2(\lambda')$ |
| | $q_2(\lambda')$ |

| **A** | **L** |
|---|---|
| $z$ | $z - \lambda$ |
| | $\lambda$ |

| **A** | **L** |
|---|---|
| $\lambda$ | $\lambda - \lambda'$ |
| | $\lambda'$ |

Figure 12: Section 3: Balance Sheets Per Unit of Wealth

| **A** | **L** |
|---|---|
| $\frac{1}{1 - q_1(\lambda)}$ | $1$ |
| | $\frac{q_1(\lambda)}{1 - q_1(\lambda)}$ |

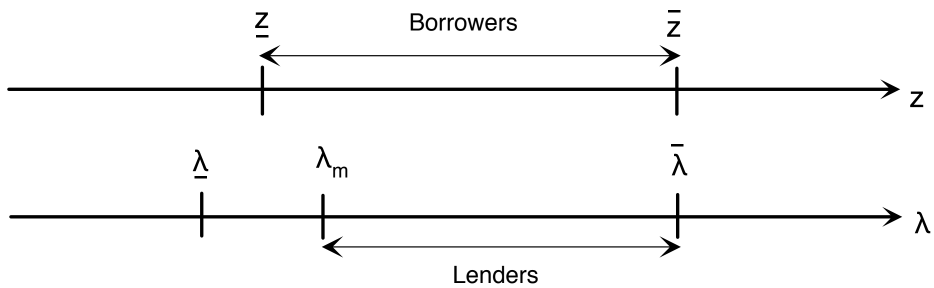
| **A** | **L** |
|---|---|
| $\frac{q_1(\lambda)}{q_1(\lambda) - q_2(\lambda')}$ | $1$ |
| | $\frac{q_2(\lambda')}{q_1(\lambda) - q_2(\lambda')}$ |

| **A** | **L** |
|---|---|
| $\frac{z}{1 - q_1(\lambda)}$ | $\frac{z - \lambda}{1 - q_1(\lambda)}$ |
| | $\frac{\lambda}{1 - q_1(\lambda)}$ |

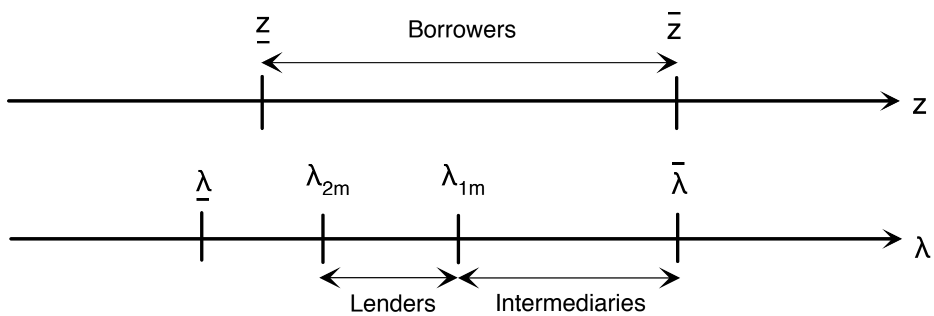
| **A** | **L** |
|---|---|
| $\frac{\lambda}{q_1(\lambda) - q_2(\lambda')}$ | $\frac{\lambda - \lambda'}{q_1(\lambda) - q_2(\lambda')}$ |
| | $\frac{\lambda'}{q_1(\lambda) - q_2(\lambda')}$ |

Figure 13: Partition of agents in different setups

(a) Setup of Section 2: simplest model




(b) Setup of Section 3: endogenous intermediary leverage




(c) Setup of Section 4: general equilibrium with occupational choice

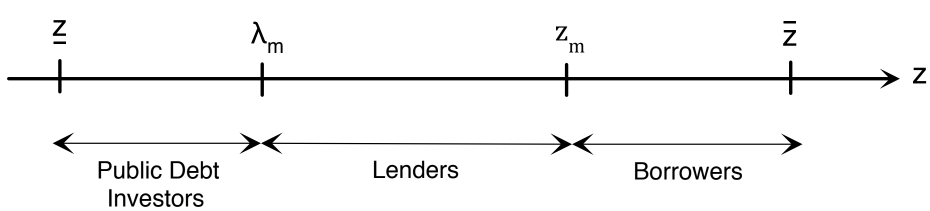
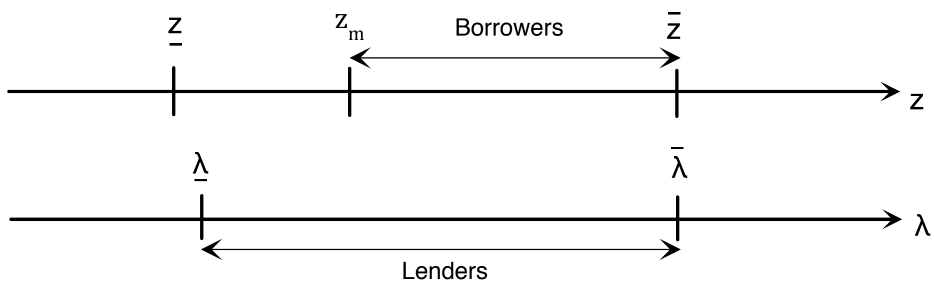
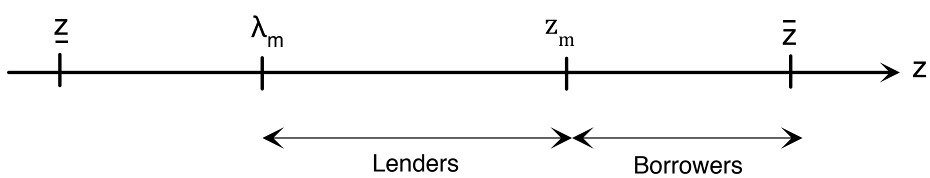


Figure 14: Partition of agents in different setups (Appendix)

(a) Setup of Appendix C: scarce lenders




(c) Setup of Appendix D: occupational choice




(d) Setup of Appendix E: uninformed depositors, regulatory capital

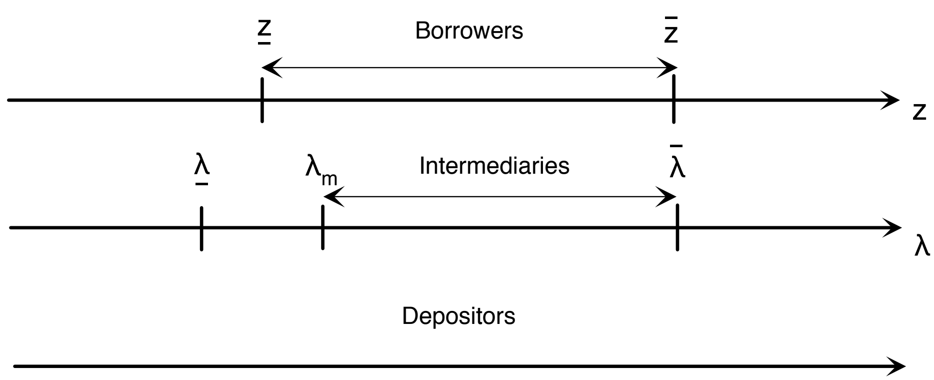


Figure 15: Intermediary-Based Asset Pricing From Adrian et al. (2014)

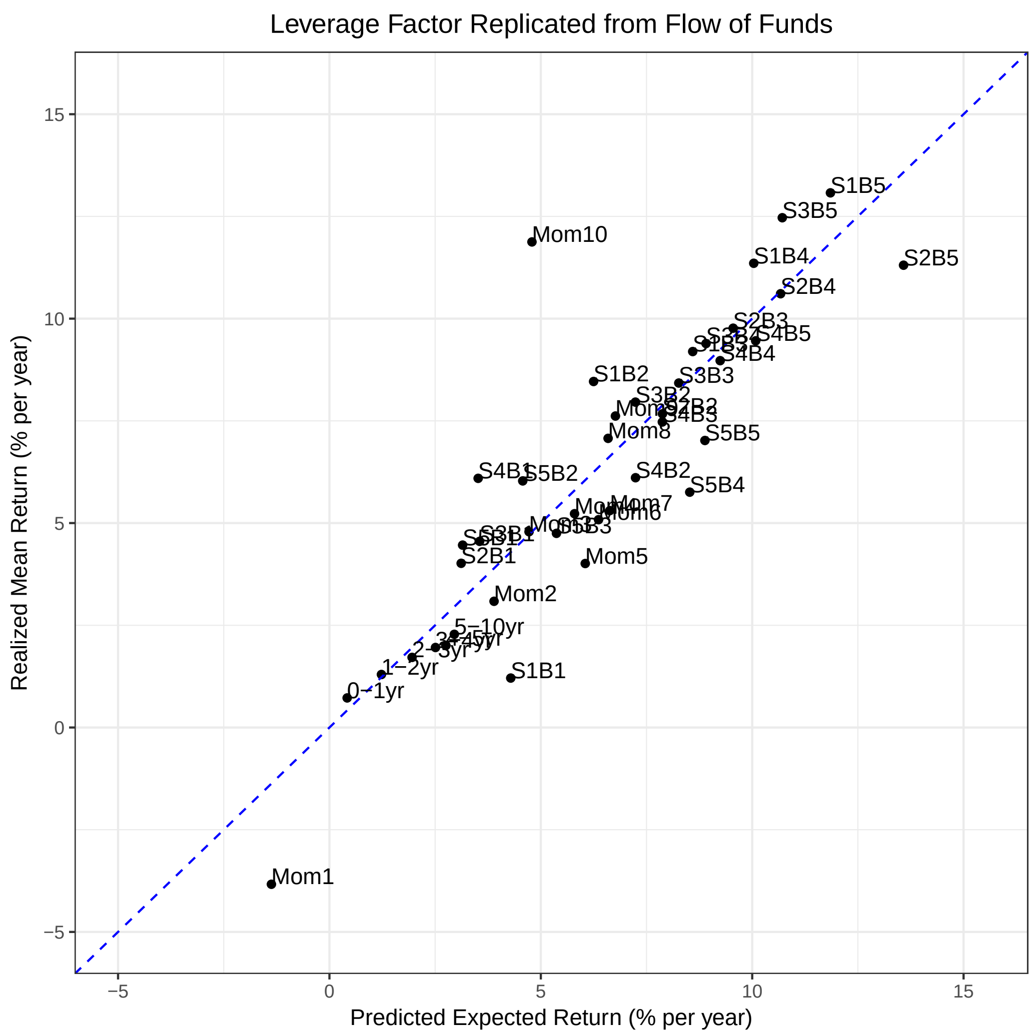


Figure 16: Expected Returns and Consumption, from Muir (2017)

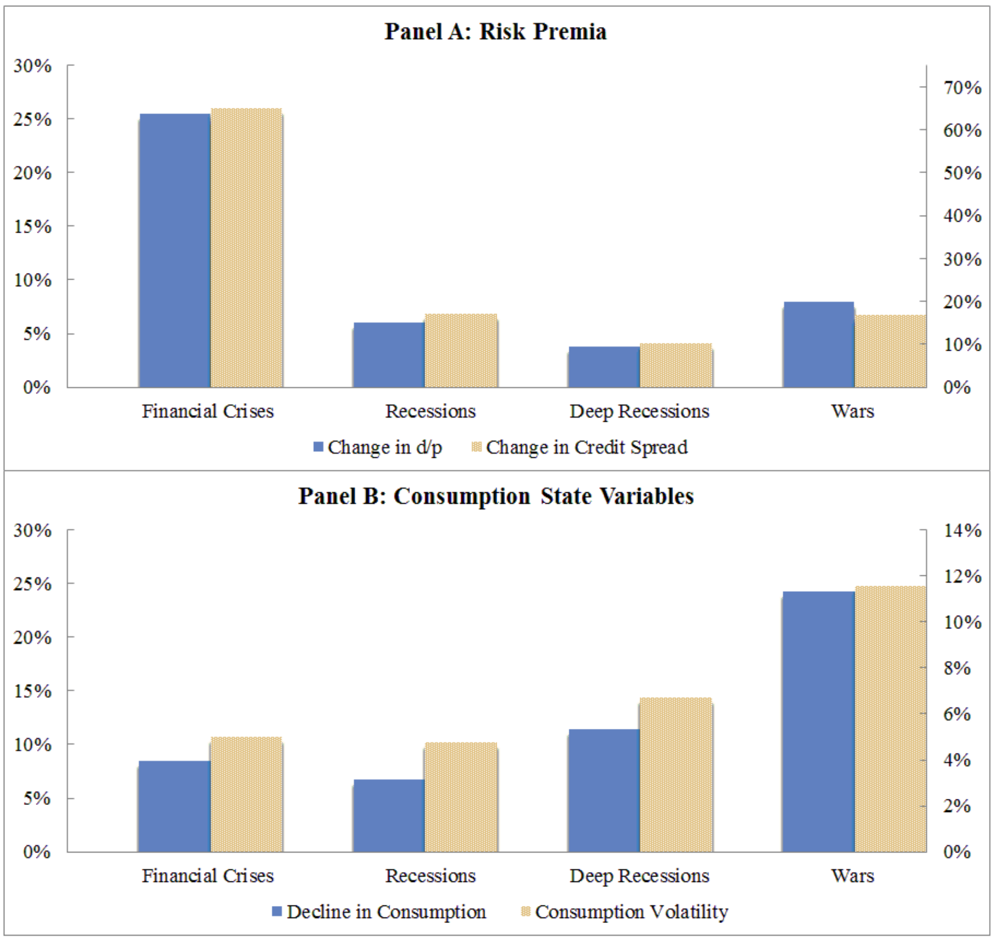


Figure 17: Replication of Krishnamurthy and Vissing-Jorgensen (2012)

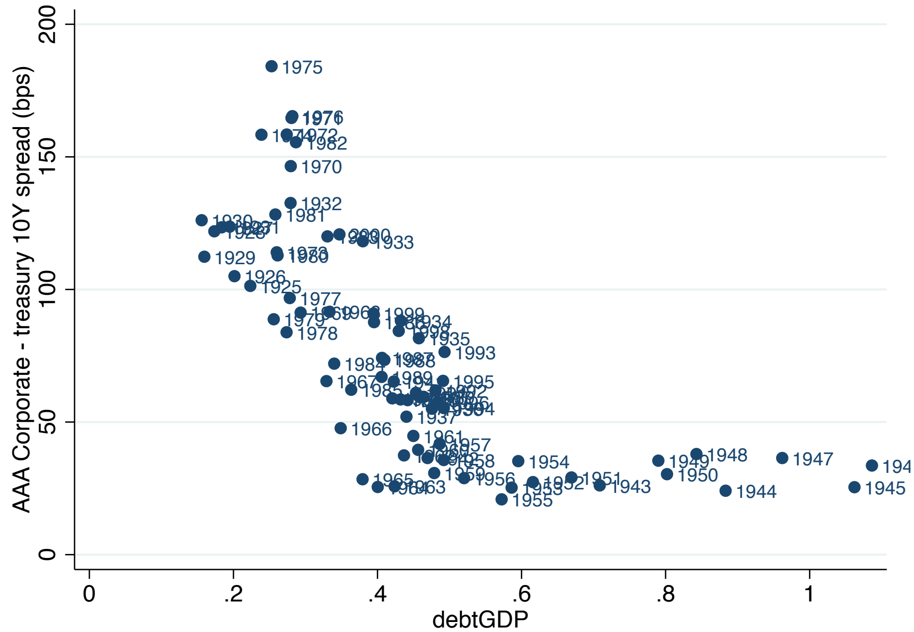



Figure 18: Replication of Krishnamurthy and Vissing-Jorgensen (2012)

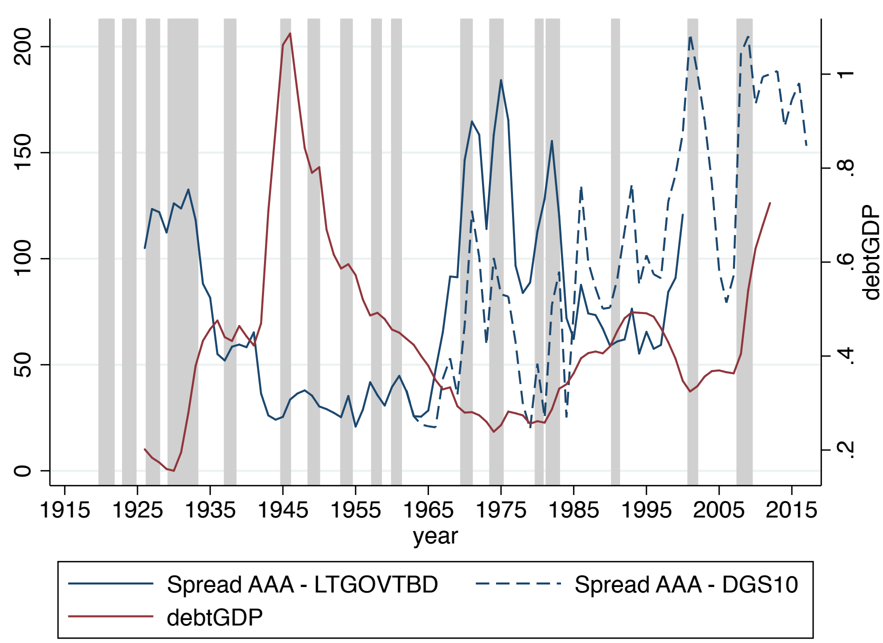


Figure 19: Corporate Bond Spreads, Inspired from Collin-Dufresne et al. (2001)

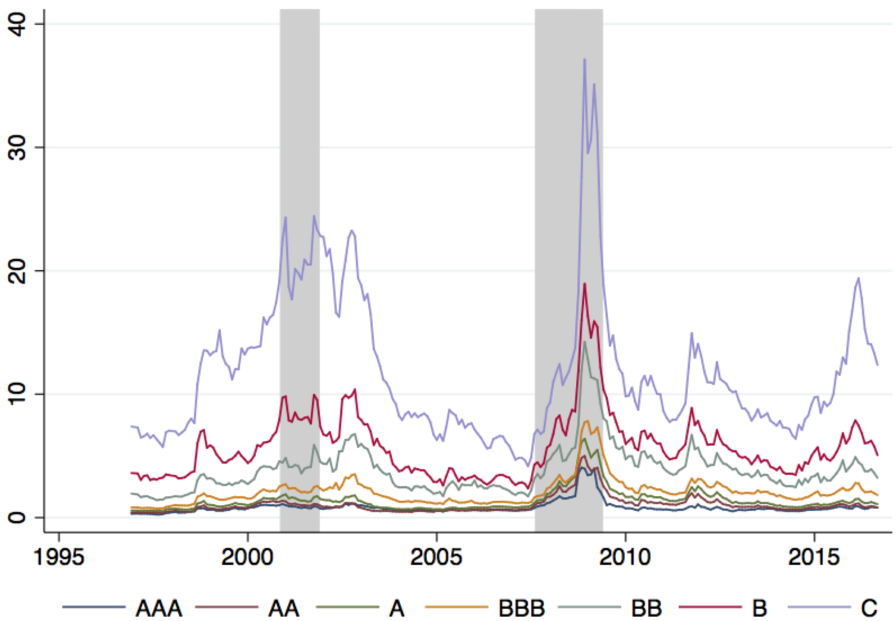



Figure 20: Log Corporate Bond Spreads, Inspired from Collin-Dufresne et al. (2001)