Stanford | Institute for Economic Policy Research (SIEPR)

The Social Value of Financial Expertise

Pablo Kurlat
Stanford University

November, 2017

Working Paper No. 17-038

The Social Value of Financial Expertise

Pablo Kurlat*
Stanford University

November 2017

Abstract

I study expertise acquisition in a model of trading under asymmetric information. I propose and implement a method to measure r, the ratio of the marginal social value to the marginal private value of expertise. This can be decomposed into three sufficient statistics: traders' average profits, the fraction of bad assets among traded assets and the elasticity of good assets traded with respect to capital inflows. I measure $r = 0.16$ for the junk bond underwriting market. Since this is less than one, it implies that marginal investments in expertise destroy surplus.

Keywords: Financial industry, expertise, asymmetric information, sufficient statistic

JEL codes: D53, D82, G14, G20

*I am grateful to Sebastian Di Tella, Bob Hall Pete Klenow, Brad Larsen, Alejandro Martins, Thomas Philippon, Monika Piazzesi, Martin Schneider, Johannes Stroebel, Victoria Vanasco and Iván Werning for helpful comments. Ricardo de la O Flores provided outstanding research assistance. I also thank Nada Mora for kindly sharing her data with me. Correspondence: Department of Economics, Stanford University. 579 Serra Mall, Stanford, CA, 94305. Email: pkurlat@stanford.edu

1 Introduction

The financial industry has been heavily criticized in recent years. One criticism often made is that it has simply become too large. Tobin (1984) worried that "we are throwing more and more of our resources, including the cream of our youth, into financial activities [...] that generate high private rewards disproportionate to their social productivity". In the decades since Tobin's remark, the financial industry has become much larger. Philippon and Reshef (2012) and Philippon (2014) document that the share of value added of financial services in GDP has risen from about 5% in 1980 to about 8% in recent years. While 8% of GDP is certainly a large number, it doesn't necessarily follow that it's excessive. In order to reach this conclusion one needs to have a framework for assessing how the size of the financial industry compares with the social optimum.

In this paper I propose and implement a method to measure a variable I label r. r is the ratio of the marginal social value to the marginal private value of dedicating resources to one activity within finance: acquiring expertise to evaluate assets. If $r > 1$, then the marginal social value exceeds the marginal private value. Under the assumption that marginal private value equals marginal cost, this implies that marginal social value exceeds marginal cost and a social planner would want more resources allocated to this activity. Conversely, if $r < 1$, a planner would want fewer resources dedicated to it.

The measurement is based on a particular model of financial expertise. I assume that financial firms earn income because they have expertise to trade in markets with asymmetric information: banks assess the creditworthiness of borrowers, venture capitalists decide which startups are worth investing in, insurance companies evaluate risks, etc. Acquiring this expertise requires using productive resources that might be employed elsewhere: talented workers develop valuation models, IT equipment processes financial data, etc.

I formalize this in a model with the following elements. There is a group of households who own heterogeneous assets, either good or bad. Each household can keep its asset or sell it to a bank. Due to differential productivity or discount factors, selling assets creates gains from trade, which differ by household. Each household is privately informed about the quality of its own asset, while banks only observe imperfect signals about them. Each bank may, at a cost, acquire expertise. Having more expertise means receiving more accurate signals about the quality of the assets on sale.

I model trading using the competitive equilibrium concept proposed by Kurlat (2016). In equilibrium, all assets trade at the same price; owners of good assets can sell as many units as they choose at that price but owners of bad assets face rationing. Only banks that are sufficiently expert choose to trade, while the rest stay out of the market. The price reflects the pool of assets acceptable to the marginal bank. Because this pool includes bad assets, households that sell good assets do so at a discount. Therefore, as in Akerlof (1970), households who have insufficiently large gains from trade choose not to sell, leading to a loss of surplus.

In this model, expertise is privately valuable to the individual bank because it enables it to better select which assets to acquire, improving returns. It is also socially valuable because it reduces overall information asymmetry, changing equilibrium prices and inducing socially valuable marginal trades. However, there is no reason for private and social values to be equal, i.e. no reason to believe $r = 1$. The private value depends on how expertise improves an individual trader's portfolio. The social value depends on how it draws marginal households into the market by shifting the entire equilibrium, which the marginal bank ignores.

It is possible to derive an analytical expression for r but it turns out to be quite complicated because it depends on various possible feedback effects. The main result I show is that it is possible to decompose the formula for r into sufficient statistics: measurable quantities that, combined, capture all the effects that are relevant for r without the need to separately measure all the parameters of the model. In particular, I show that

$$r = \eta \left(1 - \frac{1-f}{\alpha}\right) \tag{1}$$

where η is the elasticity of the volume of good assets that are traded with respect to capital inflows, f is the proportion of bad assets among the assets that are traded and α is the average NPV per dollar invested earned by banks. α and f enter formula (1) because they measure the value of marginal trades: if banks make high profits despite acquiring a high fraction of bad assets, the adverse selection discount suffered by the marginal seller must be high, indicating large gains from trade at the margin. η enters formula (1) because an inflow of funds and an increase in the expertise of an individual bank affect the equilibrium through the same channel: by increasing the demand for good assets. Therefore η is informative about how many additional trades would take place if a bank increased its expertise at the margin.

I implement formula (1) empirically for the junk bond underwriting market. I take measures of α (underwriting profitability) directly from existing studies (Datta et al. 1997, Jewell and Livingston 1998, Gande et al. 1999). I measure f by looking at the default rate of the universe of junk bonds issued between 1977 and 2010. The elasticity η is the hardest to measure, since it requires identifying plausibly exogenous variation in the capital devoted to bond underwriting. I measure this in two ways. The first is by using past profitability of investment banks as an instrument; the second is by focusing on the period around 1990, exploiting the variation around the collapse of the investment bank Drexel Burnham Lambert. Combining these measurements according to formula (1), I find $r = 0.162$. This implies that of the last dollar earned by junk bond underwriters by investing in expertise, 16 cents is value added and the rest is captured rents. By this measure underwriters overinvest in expertise.

Formula (1) applies under the polar assumptions that the value of bad assets is zero, that households know perfectly which assets are bad and that everyone is risk neutral, which are unrealistic for most applications, including junk bonds. I then explore possible alternatives to see whether the formula can be extended beyond this polar case. I first look at the possibility that bad assets have a positive value (like a recovery value for bonds in default), while retaining the assumption that payoffs are binary. Formula (1) extends readily to this case, one just needs to measure the recovery value. Next, I consider the possibility that households don't know exactly what the payoff of their asset will be. Accounting for this possibility is harder, because it requires measuring what households know, which has no immediate empirical counterpart. However, it is possible to sign the direction of the bias and establish whether the measured r is an upper or a lower bound; I find that the measured r is an upper bound. I also relax the assumption of risk neutrality and find that by using additional information on asset prices it is again possible to establish an upper bound on r. I then study the case where the distribution of asset payoffs is not binary. For this case, no simple formula exists. In order to compute r for this case, I impose more structure on the problem by making functional form assumptions and use the measured statistics as calibration targets. Finally, I allow jointly for both non-binary payoffs and less-than-perfect information by households. The estimates of r in these extensions range from 0.06 to 0.29. While there is considerable uncertainty around these numbers, the conclusion that $r < 1$ seems fairly robust.

As discussed by Cochrane (2013) and Greenwood and Scharfstein (2013), underlying some of the concern about the size of the financial industry is a view that finance is a largely rent-seeking, socially wasteful, industry. Bolton et al. (2011), Philippon (2010), Glode et al. (2012), Shakhnov (2014) and Fishman and Parker (2015) describe theoretical environments where over-investment in financial expertise emerges as an equilibrium outcome. In the model I study, banks are engaging in activities that look a lot like rent-seeking, since they dedicate resources to try to find profitable trades. However, this has the socially valuable side effect of correcting the mispricing that arises due to adverse selection, which induces gains from trade. The relative magnitude of rents and social gains could in principle go in either direction (pure rent-seeking is a special case), and can be assessed empirically.

There is a separate empirical literature on the value of finance based on aggregate cross-country data. Murphy et al. (1991) find that the proportions of university graduates in law (negatively) and engineering (positively) are correlated with economic growth, and argue that this roughly corresponds to the distinction between financial and productive activities. Levine (1997, 2005) surveys cross country evidence that finds a positive correlation between economic growth and the size of the financial sector. Relative to this literature, I take a micro rather than aggregate perspective. Instead of studying the value of finance as a whole, I focus on the marginal value of specific activities within the financial industry.

On the theoretical side, the model and the equilibrium concept are extensions of Kurlat (2016), which in turn builds on earlier ideas by Gale (1996) and Guerrieri et al. (2010). Kurlat (2016) focuses on a case where each seller either has no gains from trading or does not value the asset at all. A relatively minor innovation in the present study is to extend the analysis to the case where the gains from trade can take intermediate values. This matters because inefficient retention only happens among sellers with these intermediate potential gains from trade. The more substantial new result is the derivation of the sufficient statistic formula (1).

The paper is organized as follows. Section 2 presents the model, defines and characterizes the equilibrium and defines r. Section 3 has the derivation of the sufficient statistics needed to measure r. Section 4 presents the measurements of r for junk bond underwriting. Section 5 contains the extensions of the basic model. Section 6 discusses the implications of the findings and some of their limitations.

2 The Model

2.1 Agents, Preferences and Technology

The economy is populated by households and banks, all of whom are risk neutral.

Banks are indexed by $j \in [0, 1]$. Bank j has an endowment w_j of goods that it may use to buy assets from households. It is best to think of this endowment as including both the bank's equity and its maximum debt capacity, i.e. the maximum amount of funds it can invest.

Households are indexed by $s \in [0,1]$. Each household is endowed with a single divisible asset $i \in [0,1]$, which it may keep or sell to a bank. The household's type s and the index of its asset i are independent. If sold to a bank, asset i will produce a dividend of

$$q(i) = \mathbb{I}(i \geq \lambda) V \tag{2}$$

This means a fraction λ of assets are bad and yield 0 and a fraction $1-\lambda$ are good and yield V. If instead household s keeps asset i, it will produce a dividend of $\beta(s) q(i)$. Therefore $(1-\beta(s)) V$ are the gains produced if a household of type s sells a good asset to a bank. Assume w.l.o.g. that $\beta(\cdot)$ is weakly increasing, so higher types get more dividends out of good assets. There is no need to assume that $\beta(s) < 1$ for all s, the model can allow for households for whom there are no gains from trade.

Several applications fit this general framework. In an application to household borrowing, $q(i)$ represents future income and $\beta(s)$ is the household's discount factor. In an application to venture capital, households represent startup companies, banks represent venture capital funds and $\beta(s)$ is the fraction of the startup's potential value that can be realized without obtaining venture financing. In an application to insurance, $q(i)$ is the household's expected income net of any losses and $\beta(s) q(i)$ is its certainty-equivalent.

2.2 Information and Expertise

The household knows the index i of its asset and therefore its quality $q(i)$. Banks do not observe i directly but instead observe signals that depend on their individual expertise. A bank with expertise $\theta \in [0,1]$ will observe a signal

$$x(i,\theta) = \mathbb{I}(i \geq \lambda\theta) V \tag{3}$$

whenever it analyzes asset i, as illustrated in Figure 1.[1] Higher-θ banks are more expert because they make fewer mistakes: they are more likely to observe signals whose value coincides with the true quality of the asset.

[1]The information structure implied by equation (3) is special in that banks only make mistakes in one direction. Kurlat (2016) analyzes other possible cases.

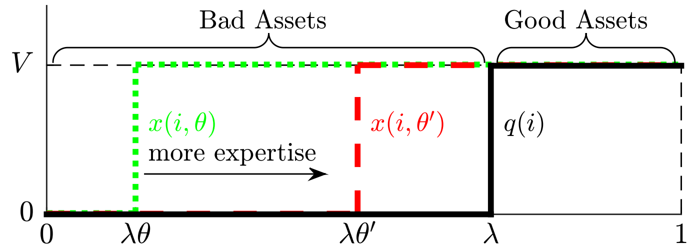

Figure 1: Asset qualities and signals

The level of expertise θ is endogenously chosen by each bank. The cost for bank j of acquiring expertise θ is given by $c_j(\theta)$. The function $c_j(\cdot)$ is allowed to be different for different banks.

2.3 Equilibrium

There are two stages. Banks acquire expertise in the first stage and trading takes place in the second. For the trading stage, I adopt the notion of competitive equilibrium proposed by Kurlat (2016). The complete definition of equilibrium and the proof that the characterization below is indeed an equilibrium are stated in Appendix A.

Markets at every possible price are assumed to coexist, and any asset can in principle be traded in any market. Households choose in what market (or markets, as there is no exclusivity) to put their asset on sale and banks choose what markets to buy assets from. Banks who want to buy may be selective, refusing to buy some of the assets that are on sale, but how selective they can be depends on their expertise. They can only discriminate between assets that their own information allows them to tell apart. This implies that a bank with expertise θ will accept assets in the range $i \in [\lambda\theta, 1]$. This range includes $i \in [\lambda\theta, \lambda)$ (some of the bad assets) and $i \in [\lambda, 1]$ (all the good assets). Banks receive a random sample of the assets on sale that they are willing to accept.

Market clearing is not imposed as part of the equilibrium definition. Assets may be offered on sale in a given market but not traded because there are not enough buyers who are willing to accept them. As in Gale (1996) and Guerrieri et al. (2010), rationing may and indeed does emerge as an equilibrium outcome.

An equilibrium in the trading stage will result in a function $\tau(\theta)$ which says what is the profit per unit of wealth of a bank with expertise θ. Given this, the first stage of the bank's problem is straightforward. Bank j chooses expertise θ_j by solving:[2]

$$\max_{\theta} w_j \tau(\theta) - c_j(\theta) \tag{4}$$

Let $W(\theta)$ denote the total wealth of banks that choose expertise at most θ, i.e.

$$W(\theta) \equiv \int w_j \mathbb{I}(\theta_j \leq \theta) \, dj \tag{5}$$

and let $w(\theta) \equiv \frac{\partial W(\theta)}{\partial \theta}$. Nothing depends on $W(\theta)$ being differentiable but it simplifies the exposition.

Taking $W(\theta)$ as given, the equilibrium in the trading stage is summarized by three objects: a single equilibrium price p^*, a marginal household s^* that is indifferent between holding or selling a good asset and a marginal expertise level θ^* that leaves the bank indifferent between buying and not buying assets.

If household s decides to retain a good asset, its payoff is $\beta(s) V$; if instead it decides to sell it, its payoff is p^*. Therefore, in equilibrium the marginal household satisfies:

$$p^* = \beta(s^*) V \tag{6}$$

The measure of households that sell good assets is s^* and the total number of good assets on sale is $(1-\lambda) s^*$.

A bank who buys at price p^* will be buying from a pool that contains $(1-\lambda) s^*$ good assets and λ bad assets, since all households that own bad assets will attempt to sell them. If it has expertise θ it will reject all assets with $i < \lambda\theta$, so the effective pool it draws from will have $\lambda(1-\theta)$ bad assets. As a result, the fraction of good assets it will buy is $\frac{s^*(1-\lambda)}{s^*(1-\lambda)+\lambda(1-\theta)}$, which is increasing in θ because more expert banks are able to filter out more bad assets. The profit per unit of wealth of a bank with expertise θ is:

$$\tau(\theta) = \frac{1}{p^*} \left[\frac{s^*(1-\lambda)}{s^*(1-\lambda) + \lambda(1-\theta)} V - p^* \right] \tag{7}$$

There is a cutoff value θ^* such that $\tau(\theta)$ is positive if and only if $\theta > \theta^*$. Rearranging leads to:

$$p^* = \frac{s^*(1-\lambda)}{s^*(1-\lambda) + \lambda(1-\theta^*)} V \tag{8}$$

[2]For simplicity, the cost $c_j(\theta)$ is expressed directly in utility terms.

Banks with expertise above θ^* spend all their wealth buying assets while banks with expertise below θ^* choose not to buy at all.

A bank with expertise θ will buy $\frac{1}{p^*}\frac{s^*(1-\lambda)}{s^*(1-\lambda)+\lambda(1-\theta)}$ good assets per unit of wealth. This means that in total, banks will buy

$$\int_{\theta^*}^{1} \frac{1}{p^*}\frac{s^*(1-\lambda)}{s^*(1-\lambda)+\lambda(1-\theta)}dW(\theta)$$

good assets. Imposing that all the $(1-\lambda)s^*$ good assets placed on sale are indeed sold and rearranging implies:

$$p^* = \int_{\theta^*}^{1} \frac{1}{s^*(1-\lambda)+\lambda(1-\theta)}dW(\theta) \tag{9}$$

Note that market clearing of good assets is a result, it's not imposed as part of the definition of equilibrium. In fact, since bad assets are rejected by at least some banks, not all the ones that are put on sale are sold in equilibrium.

An equilibrium is given by p^*, s^* and θ^* that satisfy (6), (8) and (9). Under regularity conditions stated in Appendix A, the equilibrium is unique.

2.4 Welfare

I measure welfare as the total surplus that is generated by trading assets, ignoring the distribution of gains. When a household of type s sells a good asset, this creates $(1-\beta(s))V$ social surplus. Integrating over all households that sell yields a total surplus of:

$$S = (1-\lambda)\int_{0}^{s^*}(1-\beta(s))Vds \tag{10}$$

Consider an individual bank j that in equilibrium chooses to acquire expertise θ_j. Holding the expertise choices of all other banks constant, let $S_j(\theta)$ be the social surplus that would result if instead bank j were to acquire expertise θ. Define

$$r_j \equiv \frac{S_j'(\theta_j)}{w_j\tau'(\theta_j)} \tag{11}$$

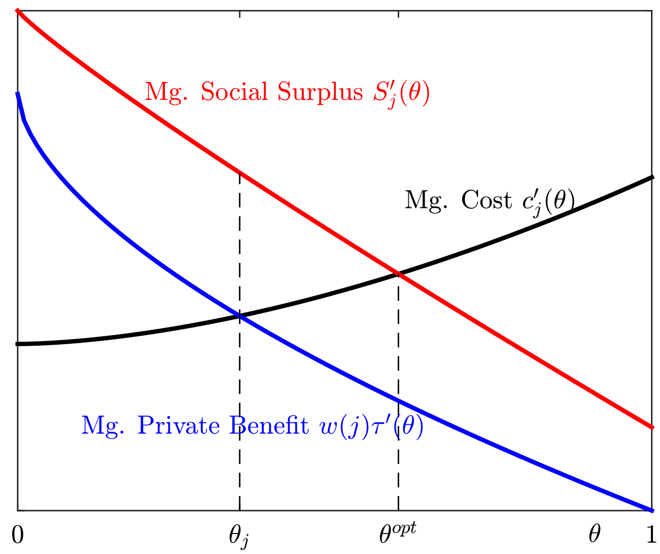

Figure 2: Example of marginal social surplus, private benefit and cost of additional investments in expertise. Bank j will choose expertise θ_j, equating marginal private benefit and marginal cost. The socially optimal level of expertise would be θ^{opt}.

Why is r_j an object of interest? The logic is illustrated in Figure 2. The first order condition for problem (4) is:

$$w_j \tau' (\theta_j) = c_j' (\theta_j)$$

and therefore

$$r_j = \frac{S_j' (\theta_j)}{c_j' (\theta_j)}$$

Hence r_j is a measure of the amount of social value created per unit of marginal resources that bank j invests in acquiring expertise. In the example in Figure 2, at the equilibrium level of expertise θ_j, we have $S_j' (\theta_j) > w_j \tau' (\theta_j)$ so $r_j > 1$, which means that at the margin investing more in expertise increases the net social surplus.

2.5 Private and Social Incentives

When a bank acquires additional expertise, it reduces the range of bad assets that it finds acceptable, improving its selection. This is the source of private incentives to acquire expertise.

Using (7), the marginal private gain from for bank j is:

$$w_j \tau' (\theta_j) = \frac{w_j}{p^*} V \frac{\lambda (1-\lambda) s^*}{\left[(1-\lambda) s^* + \lambda (1-\theta_j) \right]^2} \tag{12}$$

Changes in expertise also change the equilibrium, which affects the utility of both households and other banks. A bank's expertise affects the equilibrium through the market clearing condition (9). More expertise means that, for any given level of wealth, a bank will buy fewer bad assets and therefore more good assets. Therefore something must adjust for the market to clear. In general, all three endogenous variables will adjust. The equilibrium price p^* will rise; this will lead the marginal bank to exit, raising θ^*, and persuade the marginal household to sell assets, raising s^*.

From a social perspective, the change in price in itself is neutral: it benefits households at the expense of banks but it's just a transfer. The only thing that matters for the social surplus is the increase in s^*. Using (10), the marginal social surplus from bank j's expertise is:

$$S_j' (\theta) = (1-\lambda) (1-\beta (s^*)) V \frac{ds^*}{d\theta_j} \tag{13}$$

In equation (13), $(1-\lambda)(1-\beta(s^*)) V$ are the gains from trade that are created if a marginal household s^* decides to sell its good asset, and $\frac{ds^*}{d\theta_j}$ is the shift in s^* when bank j increases its expertise.

The basic source of inefficiency in the economy is the standard force in Akerlof (1970): households inefficiently retain good assets because by selling them into a common pool with bad assets they are unable to capture their full value. Expertise is socially valuable to the extent that it undoes this underlying inefficiency. More expert banks filter out bad assets from the pool, bid up the price and persuade marginal households to sell good assets, creating gains from trade. If expertise were free to acquire, in this economy it would always be socially beneficial to do so.

The marginal gains from trade that are created depend, among other things, on the density of households that are close to the indifference margin. Suppose for example that $\beta(s)$ was a step function of the form:

$$\beta (s) = \mathbb{I} (s \leq \mu)$$

where $\mu \in (0,1)$. It's easy to see that in this case $s^* = \mu$ for any distribution of expertise. Households with $s \leq \mu$ have no value for retaining the asset so they would always sell in equilibrium while households with $s > \mu$ value it just as much as banks so they will never sell. Since there are no households close to the indifference margin, in this case, $\frac{ds^*}{d\theta_j} = 0$ and expertise has no social value. It would still, however, have a private value because an individual bank would still benefit from better selection, so bank profits would be purely rent extraction. Conversely, if there were many households with $\beta(s)$ close to $\frac{p^*}{V}$, then a small increase in the price would induce large additional gains from trade. Since banks are small and take the equilibrium as given, the shape of $\beta(s)$ is just not part of the private cost-benefit calculation.

In a standard efficient competitive economy it's also the case that agents ignore their effect on the equilibrium, but this does not result in an inefficiency because, since all the gains from trade are exhausted in equilibrium, marginal trades create no social value. What is special about an economy with underlying information asymmetry is that the marginal trade creates strictly positive social value.

It is useful in applications to have a broad interpretation of what "selling" and "retaining" an asset means. Consider a potential entrepreneur who has a good business idea and is deciding whether to pursue it. If he can get outside funding on good terms, he will do so, effectively selling a part of his business idea to financial markets. Otherwise, he may not start a business at all and just look for a job, effectively retaining his idea and getting less out of it than the first-best use. Under this broad interpretation, the usefulness of more expert financial intermediaries is that they make it possible for good projects to be carried out, improving ex-ante investment decisions.

It is worth noting that if it were possible to redistribute banks' endowments, then investing in expertise would always be socially wasteful. Rather than having many banks invest independently in acquiring the same expertise, the efficient thing to do would be to have a single bank acquire expertise and manage everyone's endowment. The maintained assumption is that for unmodeled moral hazard or span-of-control reasons this is not possible. Studying r_j answers the question of what is the marginal social value of investments in expertise taking as given the duplicative nature of these investments.

3 Measuring r

3.1 Solving for r_j

Replacing (13) and (12) in (11):

$$r_j = \frac{(1-\lambda)(1-\beta(s^*))V}{w_j \frac{V}{p^*}\frac{\lambda(1-\lambda)s^*}{[(1-\lambda)s^*+\lambda(1-\theta)]^2}} \frac{ds^*}{d\theta_j} \tag{14}$$

A key ingredient of equation (14) is $\frac{ds^*}{d\theta_j}$, how many additional households sell good assets when the expertise of bank j changes. In order to compute this, rewrite equations (6)-(9) compactly as:

$$K(p^*, \theta^*, s^*) = 0 \tag{15}$$

where

$$K(p^*, \theta^*, s^*) = \begin{pmatrix} p^* - \beta(s^*)V \\ p^* - \frac{(1-\lambda)s^*}{(1-\lambda)s^*+\lambda(1-\theta^*)}V \\ p^* - \int_{\theta^*}^1 \frac{1}{(1-\lambda)s^*+\lambda(1-\theta)} dW(\theta) \end{pmatrix}$$

Let K_i denote the i_{th} dimension of the function K and $D = \nabla K$ denote the matrix of derivatives of K.

Using the implicit function theorem, (15) implies:

$$\frac{ds^*}{d\theta_j} = -D_{33}^{-1}\frac{\partial K_3}{\partial \theta_j} \tag{16}$$

where

$$D_{33}^{-1} = -\frac{1}{|D|}\frac{\lambda(1-\lambda)s^*}{[(1-\lambda)s^*+\lambda(1-\theta^*)]^2}V \tag{17}$$

$$|D| = \frac{V}{[(1-\lambda)s^*+\lambda(1-\theta^*)]^2}\begin{bmatrix} -\frac{\lambda(1-\lambda)(1-\theta^*)}{(1-\lambda)s^*+\lambda(1-\theta^*)}w(\theta^*) \\ +[\lambda(1-\lambda)s^*V+((1-\lambda)s^*+\lambda(1-\theta^*))w(\theta^*)]\beta'(s^*) \\ +\lambda(1-\lambda)s^*\int_{\theta^*}^1 \frac{(1-\lambda)}{[(1-\lambda)s^*+\lambda(1-\theta)]^2}dW(\theta) \end{bmatrix} \tag{18}$$

$$\frac{\partial K_3^*}{\partial \theta} = -w_j \frac{\lambda}{[(1-\lambda)s^*+\lambda(1-\theta)]^2} \tag{19}$$

Equation (19) captures the direct effect of an increase in bank j's expertise. More expertise implies rejecting more bad assets and therefore buying more good assets. This shifts the market clearing condition. Other things being equal, prices would have to rise to restore equation (9). But, of course, all the endogenous variables respond: higher prices attract marginal sellers of good assets and repel marginal banks, so both s^* and θ^* respond as well. The term D_{33}^{-1} measures how shifts in the market clearing condition translate, through all the feedback channels in the model, into a change in the marginal seller. Equation (19) implies this is always positive: more expert banks lead to a higher equilibrium price and this induces marginal households to sell good assets.

Replacing equations (16)-(19) into equation (14) and simplifying:

$$r_j = \frac{1}{|D|} \frac{\lambda (1-\lambda) (1-\beta(s^*)) p^* V}{\left[(1-\lambda) s^* + \lambda (1-\theta^*)\right]^2} \tag{20}$$

Formula (20) immediately implies the following result.

Proposition 1. *r_j does not depend on θ_j or w_j*

One might have conjectured that the misalignment of social and private returns to expertise might be different for banks with different wealth or for banks that (for instance due to different cost functions) choose different levels of θ. That turns out not to be the case. This means that if the financial industry has incentives to either over- or under-invest in expertise, this will be true across the board, and any corrective policies don't need to be applied selectively.

3.2 Sufficient Statistics

The main difficulty with quantifying expression (20) is that the expression for the determinant $|D|$ is quite complicated. This is because $|D|$ captures the magnitude of all the various feedback effects in the model: how selection depends on prices, the extensive margin of bank participation, etc. The key to the sufficient statistic approach is that it is not necessary to measure all the elements of $|D|$ separately. $|D|$ measures the strength of feedback effects with respect to any driving force; therefore it enters the formula for any elasticity that one could measure.

Let α be the average present value per dollar invested that banks obtain.[3] In the model:

$$\alpha = \frac{(1-\lambda) s^* V}{W} \tag{21}$$

[3] The model is static, so the relevant concept of profitability is the present value for a given initial investment rather than a per-period return.

where $W \equiv \int_{\theta^*}^{1} dW(\theta)$ is the total funds spent by banks who choose to trade and the numerator represents the total dividends obtained from assets acquired by banks.

Let f be the fraction of assets traded that turn out to be bad. In consumer loans, this would correspond to the default rate; in venture capital it would correspond to the fraction of ventures that fail, etc. If N is the total number of assets that are traded and G is the number of good assets that are traded, then:

$$f \equiv 1 - \frac{G}{N}$$

In the model we have:

$$N = \frac{W}{p^*} \tag{22}$$

$$G = (1-\lambda)\, s^* \tag{23}$$

and therefore:

$$f = 1 - \frac{(1-\lambda)\, s^* p^*}{W} \tag{24}$$

Notice that measuring f only requires tracking failures among assets that actually trade, not among all projects, which would be harder to measure. It is not necessary, for instance, to measure counterfactual default rates among applicants that are denied credit.

Using (24), (21) and (6) results in:

$$\frac{1-f}{\alpha} = \frac{p^*}{V} = \beta(s^*) \tag{25}$$

Formula (25) implies that measuring α and f makes it possible to recover $\beta(s^*)$, the value of the asset to the marginal seller, and therefore the social gains from the marginal trade. The formula has the following interpretation. If $\frac{1-f}{\alpha}$ is low, this means that banks obtain high profits despite the fact that only a small fraction of the assets they buy are good. For this to be true it must be that $\frac{p^*}{V}$ is low, i.e. they must be making very high profits on the good assets that they do buy, which means that the marginal household s^* is preventing large gains from trade by not selling.

The fact that $1 - \frac{1-f}{\alpha}$ measures the social value of the marginal trade seems likely to be relatively robust to variations in the assumptions of the model. It is simply a way to measure the adverse selection discount that sellers of good assets bear. For instance, suppose that instead of having competitive banks there was a monopsony. The monopsonist would set a lower price, so marginal sellers would choose to retain their asset; the shift in the marginal seller would imply a higher social value of the marginal trade, and this would be captured accurately by the higher α that the monopsonist is able to obtain. On the other hand, the fact that there is a single marginal trade relies on the fact that $q(i)$ is binary; otherwise there would be a different social value of the marginal trade for each i. It would still be the case that for each i the discount borne by sellers (and therefore the social value of the marginal trade) would be related to the lower-quality assets that the asset is pooled with in equilibrium.

Suppose now that there is an exogenous capital inflow into banks that increases all banks' endowments by Δ, from w_j to $(1+\Delta)\,w_j$. For instance, this could be the result of a relaxation in leverage limits that lets banks manage larger portfolios with the same net worth. According to the model, the elasticity of G with respect to this increase is

$$\begin{aligned}\eta &\equiv \frac{d\log(G)}{d\Delta} \\ &= \frac{d\log(s^*)}{d\Delta} \\ &= -D_{33}^{-1}\frac{\partial K_3}{\partial\Delta}\frac{1}{s^*} \\ &= \frac{1}{|D|}\frac{\lambda(1-\lambda)}{[(1-\lambda)s^*+\lambda(1-\theta^*)]^2}p^*V\end{aligned} \tag{26}$$

Replacing (21), (24) and (26) into (20) and rearranging results in equation (1):

$$r = \eta\left(1-\frac{1-f}{\alpha}\right)$$

η enters the formula because it is a way to measure the strength of the extensive margin $\frac{ds^*}{d\theta}$, relative to the size of private gains. An increase in the expertise of one bank affects the equilibrium through the same channel than an inflow of funds for all banks: through the market clearing condition (9). An inflow of funds means that the more expert banks can afford to buy more assets; prices must rise to restore equilibrium and s^* responds to this. An increase in expertise means that the same bank will reject more bad assets and therefore buy more good ones. Again, prices must rise to restore equilibrium and s^* responds. Both effects involve the same mechanism and the same feedback channels. Furthermore, the magnitude of the effect of expertise on the market clearing condition, given by equation (19), depends on the extent to which additional expertise allows the bank to filter out bad assets. This is the same force that determines the magnitude of the marginal private value of expertise. Hence $\tau'(\theta_j)$ cancels out from the denominator of (11).

Expertise is socially valuable because by shifting the equilibrium it induces marginal households to sell good assets. This value is large when there is a large mass of good assets near the indifference margin. But this is exactly what would lead to large η, since those same good assets would shift into the market in response to the price change that results from a capital inflow. The critical assumption behind this is that banks' ability to deploy their expertise is proportional to their capital. If one were to break that connection, for instance by assuming that each bank can only purchase a fixed number of assets, independent of its wealth, then η would no longer be informative about the effects of additional expertise.

3.3 Measuring the sufficient statistics

The quantities α and f can be measured relatively straightforwardly because they are simple averages. An alternative is to measure $\frac{p^*}{V}$ directly and use equation (25) to sidestep the need to measure α and f. Depending on the application, one approach may be more practical than the other; the empirical counterpart of V may not always be clear.

The elasticity η is more challenging because it requires identifying a plausibly exogenous capital inflow or outflow and measuring its consequences. If such identifying assumptions are satisfied, there are a few different ways to measure η depending on what outcomes are easier to measure. The first, if the number of good assets traded can be measured, is simply to measure $\eta = \frac{d\log(G)}{d\Delta}$ directly. The second is almost as simple: if one can measure total number of assets traded and failure rates, then relying on (24) one gets:

$$\eta = \frac{d\log(1-f)}{d\Delta} + \frac{d\log N}{d\Delta} \tag{27}$$

A third option, if one measures failure rates, prices and total funds invested, is to use (22) to further decompose:

$$\eta = \frac{d\log(1-f)}{d\Delta} + \frac{d\log(W)}{d\Delta} - \frac{d\log(p^*)}{d\Delta} \tag{28}$$

If a capital inflow is followed by a large improvement in quality (rise in $1-f$) this means that more good assets are being traded, so η is high. Conversely, if it is followed by a large price increase, it means that the total number of assets being traded has not increased as much, so η is low.

In all cases, measuring elasticities with respect to Δ requires measuring Δ itself, i.e. how much banks' endowments change. In some cases it might be possible to do this directly, for instance if there is an increase in leverage limits that expands maximum balance sheets by a known factor. In other cases one might have to rely on measured changes in the the total number of funds actually invested in buying assets, which is not exactly the same. One of the things that can happen when Δ increases is that, because prices rise, marginal banks exit. Therefore the measured proportional change in total funds spent buying assets could be an underestimate of Δ. Formally:

$$\frac{d\log\left(W\right)}{d\Delta} = 1 - \frac{\frac{d\theta^*}{d\Delta} w\left(\theta^*\right)}{W} \leq 1$$

However, it is not unreasonable to assume that $w\left(\theta^*\right) = 0$. Choosing $\theta = \theta^*$ means that a bank would earn $\tau\left(\theta^*\right) = 0$ despite having invested a strictly positive amount of resources in acquiring expertise. Assuming $w\left(\theta^*\right) = 0$ means assuming that no banks choose to do this. Under this assumption, measuring an elasticity with respect to measured capital flows and with respect to Δ is equivalent, i.e.

$$\frac{d\log\left(W\right)}{d\Delta} = 1 \tag{29}$$

and therefore $d\Delta$ can be replaced with $d\log\left(W\right)$ in formulas (27) or (28).

4 Application to Junk Bond Underwriting

I map the junk bond market to the model following the "certification" view of underwriting proposed by Booth and Smith (1986). The companies issuing bonds correspond to the households in the model, investment banks that underwrite bonds correspond to the banks in the model and the assets are streams of cashflows.

I abstract from the institutional and contractual complexities of underwriting and assume that it takes place as follows. Underwriters compete to buy bonds in a market that operates as described in Section 2, so that underwriter θ buys bonds in the range $i \in [\lambda\theta, 1]$ at price p^*. It can then, through some reputational mechanism, credibly disclose that the bonds it purchased are in the interval $[\lambda\theta, 1]$ and resell them to completely uninformed outside investors at a price $p\left(\theta\right) = \frac{s^*(1-\lambda)}{s^*(1-\lambda)+\lambda(1-\theta)} V$, which gives outside investors zero profits. This means that underwriter θ earns the profits indicated by equation (7); it buys bonds at p^*, which is the fair price conditional on the information of the marginal underwriter θ^* but re-sells them at $p(\theta)$, which is a fair price conditional on its own, better, information. This is equivalent to having the underwriter charge the bond issuer a fee $p(\theta) - p^*$ to certify that the bond indeed lies in the interval $[\lambda\theta, 1]$, allowing the bond issuer to sell the bond for $p(\theta)$ instead of p^*. I therefore map the model concept of bank profits to underwriting spreads.

Gande et al. (1999) report underwriting spreads averaging 2.76% for bonds rated between Caa and Ba3 between 1985 and 1996; Jewell and Livingston (1998) report similar figures. Furthermore, Datta et al. (1997) report an average initial-day return of 1.86% for low-grade bonds. Arguably, this is also part of the underwriter's compensation since it allows the underwriter to place the bonds with favored clients or bolster its reputation. Accordingly, I add these two fees to make up a total underwriting spread of 4.62% and set $\alpha = 1.046$. To the extent that underwriting spreads represent compensation for other services besides certification, such as contacting investors, this figure is an upper bound on α, and using it yields an upper bound on r.

I use data from the Securities Data Company, Bloomberg and the NYU Salomon Center to construct a database on the universe of junk bonds issued between 1977 and 2010 . For each bond, I know (up to some missing entries) the date of issuance, total dollar amount, price at issuance, coupon rate, maturity and a binary indicator of whether it subsequently defaulted. I use this the compute a standardized price for each bond (see Appendix C for details). For each year t, I define: p_t as the dollar-weighted median price of bonds issued in that year, f_t as the dollar-weighted fraction of bonds issued in that year that subsequently defaulted, and W_t as the total dollar amount of bonds issued in that year.

I then estimate the following relations:

$$\begin{aligned} \log p_t &= \alpha_p + \gamma_p \log (W_t) + \delta_p X_t + \epsilon_{pt} \\ \log (1 - f_t) &= \alpha_f + \gamma_f \log (W_t) + \delta_f X_t + \epsilon_{ft} \end{aligned} \tag{30}$$

where X_t are control variables. Applying formula (28) and assumption (29), the parameter of interest is $\eta = 1 + \gamma_f - \gamma_p$.

The problem with estimating (30) by OLS is that W_t is endogenous, which can lead to biased estimates. For instance, if a general improvement in the creditworthiness of junk bond issuers (in terms of the model, a fall in λ) attracts underwriters to enter the market, this will lead to a positive association between W_t and $(1 - f_t)$, leading to an overestimate of γ_f. A similar type of mechanism would lead to an overestimate of γ_p, so the net bias in the estimate of η is unclear.

I address this possible bias in two ways. The first is by using past profitability (measured as earnings/net worth) of investment banks as an instrument for $\log(W_t)$. The idea is that past profits increase the capital available to investment banks, including their underwriting divisions, which can then increase the volume of junk bonds they underwrite. The identifying assumption is that bank profitability at time $t-1$ affects new junk bond prices at time t and the fraction of subsequent defaults among the time-t cohort only by changing the amount of capital dedicated to junk bond underwriting. The most likely reason that this assumption would be violated is if a general change in economic conditions affects both bank profitability and the junk bond market. To guard against this, I include GDP growth at $t-1$, the return on the S&P 500 index during period $t-1$ and the cyclically adjusted price-earnings ratio on the stock market at the end of period $t-1$ as additional controls. This isolates the variation in junk bond issuance volume that is explained by past investment-bank-specific profitability and not by general changes in macroeconomic variables or asset prices. The results are reported on Table 1. In all cases I estimate by seemingly unrelated regressions to allow for correlation between ϵ_{pt} and ϵ_{ft}; this just affects the standard error of $\hat{\eta}$.

Table 1: Elasticity of junk bond prices and default rates with respect to volume - full sample.

	SURE - OLS		SURE - OLS		SURE - IV	
Dep. variable	$\log p_t$	$\log(1-f_t)$	$\log p_t$	$\log(1-f_t)$	$\log p_t$	$\log(1-f_t)$
$\log(W_t)$	0.0517	0.0380	0.0512	0.0537	0.0230	0.0415
	(0.0106)	(0.0103)	(0.0119)	(0.0096)	(0.0274)	(0.0211)
implied η	0.986		1.002		1.019	
	(0.0145)		(0.0139)		(0.0303)	
time trend	Y	Y	Y	Y	Y	Y
controls	N	N	Y	Y	Y	Y
N	34	34	34	34	34	34

First stage	
Dep. variable	$\log(W_t)$
I.B. profit	7.189
	(2.567)
time trend	Y
controls	Y
N	34
F-stat	7.56

The first stage coefficient implies that a 1 percentage point increase in investment banks' return on equity (from an average of 13%) in year t leads, other things being equal, to a 7.2% increase in the total volume of junk bonds issued in the following year, an economically and statistically significant effect (p-value 0.0046), which makes estimation by instrumental variables possible. The IV coefficient estimates imply $\hat{\eta} = 1.019$, with a 95% confidence interval of $[0.959, 1.078]$. As expected, $\hat{\gamma_p}$ and $\hat{\gamma_f}$ are both positive. A capital increase in investment banks leads to an increase in the price of junk bonds and at the same time to an increase in the effective expertise available for underwriting them, leading to an improvement in the quality of bonds issued and lower subsequent defaults. The elasticities are both relatively small and of similar magnitude, which leads to estimates of η relatively close to 1. The OLS estimates $\hat{\gamma_p}$ and $\hat{\gamma_f}$ are slightly larger than the IV estimates, which is what one would expect if volume responds endogenously to variations in issuers' creditworthiness. Nevertheless, the difference is small, especially relative to the standard errors.

The second way I address the possible bias is is to focus more narrowly on the period around 1990 in order to exploit the SEC investigations that led to the bankruptcy of the investment bank Drexel Burnham Lambert as a source of variation. Drexel was a major participant in the junk bond market, with a market share above 40%. In 1985 the SEC began investigating a series of insider-trading activities that eventually led to a number of Drexel bankers being prosecuted, and the firm itself filed for bankruptcy in February 1990.[4] Drexel's troubles were, arguably, the major driver of variation in the junk bond market in the late 1980s and early 1990s (Brewer and Jackson 2000), although the recession of the early 1990-1991 recession makes the issue not completely clear-cut. I exploit this variation to run regressions (30) by OLS for the period 1986-1992, which includes the main developments in the SEC investigation, Drexel's bankruptcy and its immediate aftermath. Figure 3 shows issuance volume during this period, and the results are reported on Table 2.[5] The sample period contains just seven observations, so the estimates are imprecise, but they are broadly in agreement with those from Table 1. The bankruptcy of Drexel meant that the expertise of its bankers was (temporarily, since over time many went to work for other banks) not combined with capital, so it was effectively absent from the market, leading to lower volume, lower prices and more subsequent defaults among those cohorts. The implied elasticity is $\hat{\eta} = 0.964$.

[4]See Stewart (1992) for a timeline of the Drexel investigation.

[5]Changing the estimation window does not make much difference.

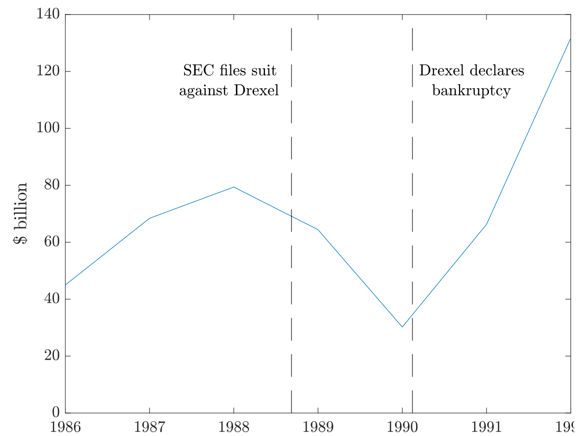

Figure 3: Yearly volume of junk bonds issued, 1986-1992.

Table 2: Elasticity of junk bond prices and default rates with respect to volume - 1986-1992 sample.

	SURE - OLS	
Dep. variable	$\log p_t$	$\log(1-f_t)$
$\log(W_t)$	0.0566 (0.0431)	0.0208 (0.011)
implied η	0.964 (0.047)	
time trend	N	N
controls	N	N
N	7	7

The dollar-weighted fraction of bonds that default across the entire sample is 0.12. This may be downward biased due to truncation, since many of the bonds issued in the later part of the sample are still outstanding and may default later on. In practice, ending the sample earlier does not make much difference. Accordingly, I set $f = 0.12$.

Replacing the point estimates $\alpha = 1.046$, $\eta = 1.019$ and $f = 0.12$ into formula (1) gives $r = 0.162$. This means that out of the last dollar that junk bond underwriters earn by being good at certifying the quality of bond issuers, 16 cents are value added and the remainder is captured rents. The 95% confidence interval is $r \in [0.153, 0.171]$.

The main reason for the low value of r is that applying formula (36) gives $\beta(s^*) = 0.841$. This is a result of the low measured values of α and f. High-quality junk bond issuers are suffering an adverse selection discount, but it's not very large. This means that the trades that produce large social gains do take place in equilibrium, and the social gains from the marginal trade are not very large, only 16% of the value of good assets.

Figure (4) shows how r changes with each of α, f and η. r is especially sensitive to default rates, but even using quite higher numbers for f leaves r comfortably below 1. Therefore the conclusion that there is more expertise acquisition than in the social optimum is fairly robust.

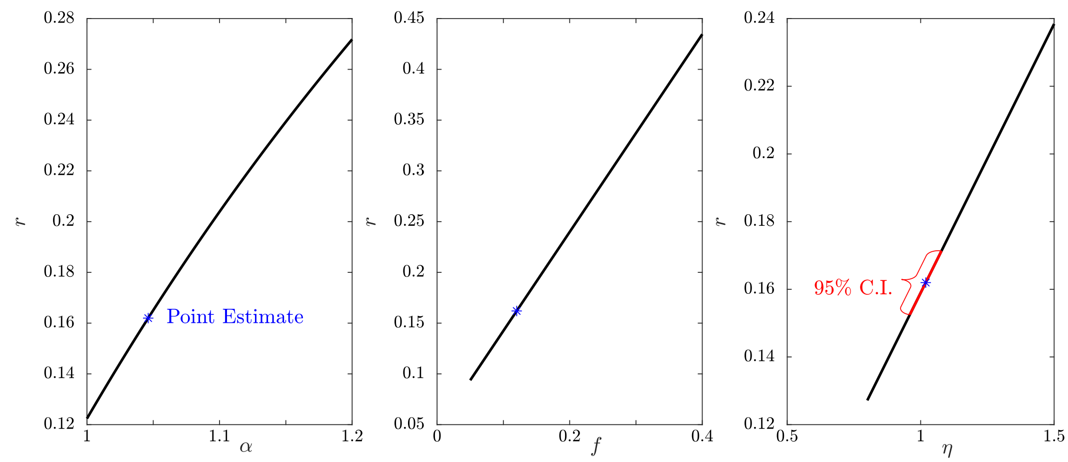

Figure 4: Sensitivity of r to each variable.

5 Relaxing some of the Assumptions

Formula (1) applies under admittedly extreme assumptions: assets' payoffs are binary, households know exactly what the dividend from their asset will be, the dividend from bad assets is exactly zero, and everyone is risk neutral. While useful for theoretical clarity, these assumptions are clearly unrealistic for most applications, including the junk bond market. In this section, I ask to what extent one can still say something about r while relaxing these assumptions.[6]

5.1 Recovery Value

Maintain the assumption that payoffs are binary and households know the quality of their asset exactly, but suppose that instead of zero, the dividend from bad assets is ϕV, where $\phi \in (0, 1)$. Under this assumption, there are two markets (prices) with active trade in

[6]One could imagine many other extensions of the model: noncompetitive behavior by banks, allowing for dynamics, signaling, etc. How these possibilities would affect the results is an open question.

equilibrium. One of them is similar to what happens in the baseline model: both good and bad assets are on sale and only sufficiently expert banks buy at this price. The other is the $p = \phi V$ market, where only bad assets are on sale. In this market, less-expert banks are willing to buy any asset on sale and make zero profits. Equations (6) and (9) still apply, while the marginal bank indifference condition (8) generalizes to:

$$p^* = \frac{s^* (1-\lambda) + \lambda (1-\theta^*) \phi}{s^* (1-\lambda) + \lambda (1-\theta^*)} V \tag{31}$$

p^* is increasing in ϕ because, other things being equal, banks are willing to pay more for an asset that has a positive recovery value.

Assume that $\beta(s) \leq 1$ for all s, so that households always value assets less than banks. The total social surplus is now:

$$S = \left[(1-\lambda) \int_0^{s^*} (1-\beta(s))\, ds + \lambda \phi \int_0^{1} (1-\beta(s))\, ds \right] V \tag{32}$$

The first term in (32) is the same as in the baseline case. The second term measures the gains from trade from bad assets. This second term does not depend on any endogenous variables. The reason is that all the bad assets that don't trade at p^* will end up trading at $p = \phi V$ instead. Changes in the equilibrium will affect at what price they trade but not whether they trade or not.[7] Using (32), the expression for the marginal social surplus is unchanged, still given by (13). The marginal private value of expertise is now:

$$w_j \tau'(\theta_j) = \frac{w_j}{p^*} V \frac{\lambda (1-\lambda) s^* (1-\phi)}{\left[(1-\lambda) s^* + \lambda (1-\theta_j)\right]^2} \tag{33}$$

Other things being equal, a higher recovery value lowers the marginal private return to expertise, because banks have less to lose from buying a bad asset. Following the same steps as in the baseline model, equation (17) generalizes to:

$$D_{33}^{-1} = -\frac{1}{|D|} \frac{\lambda (1-\lambda) s^* (1-\phi)}{\left[(1-\lambda) s^* + \lambda (1-\theta^*)\right]^2} V \tag{34}$$

[7] If $\beta(s) > 1$ for some s this would no longer be true because a household with $\beta(s) \in \left(1, \frac{p^*}{\phi V}\right)$ would put a bad asset on sale at p^* (where it would create a social *loss* if it trades) but would not sell it at $p = \phi V$. Therefore the (net) gains from trades of bad assets would be endogenous. I consider this possibility below.

Other things being equal, a higher recovery value means that s^* responds less strongly to changes in the market clearing condition, because the recovery value makes the price respond less strongly. The terms $1-\phi$ in equations (33) and (34) cancel out and expression (20) for r_j remains unchanged.

Despite the fact that the formula for r does not change, the recovery value does enter the expressions for the sufficient statistics. Formulas (22)-(24) for N and f do not change.[8] The formula for α takes into account the dividends banks obtain from bad assets and therefore becomes:

$$\alpha = \frac{N\left[(1-f)+f\phi\right]}{\int_{\theta^*}^{1} dW(\theta)} \tag{35}$$

Formula (25) generalizes to

$$\frac{1-f+f\phi}{\alpha} = \beta(s^*) \tag{36}$$

and formula (26) for η becomes:

$$\eta = -\frac{1}{|D|}\frac{\lambda(1-\lambda)(1-\phi)}{\left[(1-\lambda)s^*+\lambda(1-\theta^*)\right]^2}Vp^* \tag{37}$$

Replacing (22), (24), (35) and (37) into (20) results in:

$$r = \frac{\eta}{\alpha}\left(\frac{\alpha-1}{(1-\phi)}+f\right) \tag{38}$$

which reduces to equation (1) for the special case of $\phi = 0$.

Therefore the general approach to measuring r remains possible. It just requires measuring recovery values for bad assets. Higher measured recovery values will result in higher measured r because higher recovery values lessen private incentives for acquiring expertise without changing its social value.

Several studies have documented recovery values for bonds that default. Altman (1992) reports a average 41% recovery rate for high yield bonds between 1985-1991; Altman et al. (2005) report an average of 37.2% for all defaulted corporate bonds between 1982 and 2001; Reilly et al. (2009) report an average 42% for high yield bonds between 1987 and 2009. Mora (2015) reports an average of 39% for all bonds between 1970 and 2008. Based on these studies, I replace $\phi = 0.4$ in equation (38); together with the measured values of α, f and η, this implies $r = 0.192$. with a 95% confidence interval of $r \in [0.1807, 0.2031]$.

[8] This assumes that the data from which N and f are computed is drawn from the p^* market only and not from the $p = \phi V$ market.

5.2 Failures of Good Assets

Another stark assumption in the baseline model is that households know the dividend of their project perfectly, so that whenever an asset fails the household knew ex-ante that this was going to happen. Under this assumption, the recovery rate from assets that fail (which is what one would measure) is also the value of bad assets relative to good ones (which is what matters in the model). Also, the measured rate of failed assets corresponds exactly to the fraction of traded assets that are known to be bad at the time they are sold.

Suppose instead that even good assets fail with probability π. Let $\hat{\phi}$ denote the measured recovery rate from assets that fail. The value of bad assets relative to good ones is:

$$\phi = \frac{\hat{\phi}}{\pi\hat{\phi} + (1-\pi)} \tag{39}$$

Let $\hat{f}$ denoted the measured fraction of assets that fail. If f is the fraction of traded assets that are bad, then:

$$\hat{f} = f + (1-f)\,\pi$$

or, rearranging:

$$f = 1 - \frac{1-\hat{f}}{1-\pi} \tag{40}$$

Replacing (39) and (40) into (38) to express r in terms of measurable quantities yields:

$$r = \frac{\eta}{\alpha}\left((\alpha-1)\left(1 + \frac{\hat{\phi}}{(1-\pi)\left(1-\hat{\phi}\right)}\right) + 1 - \frac{1-\hat{f}}{1-\pi}\right) \tag{41}$$

which reduces to (38) for the special case of $\pi = 0$.

In general, given measures of α, η $\hat{\phi}$ and $\hat{f}$, r can be increasing or decreasing in π. A positive value of π means that $\hat{\phi}$ is an underestimate of ϕ: the expected value of good and bad assets is not as far apart as simply measuring recovery rates would suggest, since some good assets also fail. Since r is increasing in ϕ, correcting this bias would result in a higher measured r. On the other hand, a positive value of π means that $\hat{f}$ is an overestimate of f: measured failure rates include some good assets. Since r is increasing in f, correcting this bias would result in a lower measured r.

Obtaining a direct empirical measurement of π is difficult because π measures the extent to which households are better informed that banks ($\pi = 0$ is the polar case of pure information asymmetry), which is hard for an econometrician to observe. However, it is possible to assess which way the net bias goes without assigning an precise value to π. Taking the derivative of (41):

$$\frac{\partial r}{\partial \pi} = \frac{\eta}{\alpha \left(1-\pi\right)^2} \left(\left(\alpha - 1\right) \frac{\hat{\phi}}{1-\hat{\phi}} - \left(1 - \hat{f}\right) \right)$$

This implies that if:

$$\Upsilon \equiv \left(\alpha - 1\right) \frac{\hat{\phi}}{1-\hat{\phi}} - \left(1 - \hat{f}\right) > 0 \tag{42}$$

then the underestimate of ϕ is more severe than the overestimate of f and the value that results from assuming $\pi = 0$ and applying (38) is a lower bound for r, and vice-versa.

Replacing the measured values of α, $\hat{\phi}$ and $\hat{f}$ in (42) gives $\Upsilon = -0.849$ and since this is below zero, this implies that $r = 0.192$ is an upper bound: allowing for the possibility that some defaults arise from good assets being risky rather than pure information asymmetry would only lower r, and strengthen the conclusion that the junk bond underwriting industry dedicates too many resources to the acquisition of expertise.

5.3 Risk Aversion

The baseline model assumes that all agents are risk neutral. Suppose instead they are risk averse, and asset payoffs are exposed to aggregate risks. This exposure can arise in more than one way, with different consequences for the calculation of r.

Take the baseline case first, returning to the assumption that bad assets pay zero, but assume that the payoff from good assets is random. Under this assumption, the correct interpretation of V is the expected payoff from a good asset under banks' risk-neutral measure. In terms of the model, this is inconsequential, but risk must be taken into account when measuring α. Average profits should be computed using the risk neutral measure; using the physical measure would wrongly count as a return to expertise what is actually compensation for bearing aggregate risk. In the junk bond application, measuring α through underwriting spreads cleanly isolates the profits that underwriters obtain at the time of issuance from any excess returns that bond investors subsequently obtain, which is just a risk premium. Therefore under the baseline assumptions, no risk adjustment needs to be made when applying formula (1).[9]

[9] As long as the risk premium is constant, there should be no bias in the estimate of the elasticity of the price to capital inflows. Of course, risk premia do vary, which is one reason to control for proxies of this in estimating (30).

Next suppose that bad assets have a positive but risky recovery value. The correct empirical counterpart of ϕ in formula (38) is the expected recovery rate under the risk-neutral measure. Recovery rates are positively correlated with macroeconomic and financial conditions (Mora 2015), so risk-neutral expected recovery rates are lower than measured averages. Adjusting ϕ downwards in formula (38) would lead to a lower value for r, strengthening the conclusion that $r < 1$.

Finally, suppose that recovery rates are constant but the probability of failure from good assets is positive and risky. Formula (39) for the relative value of bad assets still applies, replacing π with π^R, the risk-neutral expected probability of failure of good assets. Empirically, default rates are negatively correlated with macroeconomic and financial conditions (Giesecke et al. 2011), so π^R is higher that the physical probability of failure of good assets. Adjusting π upwards implies higher ϕ and therefore a higher value for r.

It is still possible to find an upper bound for r by using direct information on bond prices. Normalizing the face value of bonds to 1,[10] expected profits satisfy:

$$\alpha = \frac{(1-f)\left(\pi^R\hat{\phi} + \left(1-\pi^R\right)\right) + f\hat{\phi}}{p^*} \tag{43}$$

The numerator is the expected dividend from an asset (which is good with probability $1-f$) under the risk-neutral measure.

The fraction of bad assets f is still related to measured defaults $\hat{f}$ by (40), where π is the physical probability of good assets defaulting. Replacing (40) in (43) and solving for π^R results in:

$$\pi^R = 1 - \frac{1-\pi}{1-\hat{\phi}}\frac{p^*\alpha - \hat{\phi}}{1-\hat{f}} \tag{44}$$

Equation (44) describes a negative relationship between the price p^* and the risk-neutral probability of good assets defaulting, other things being equal.[11] Given a certain observed default rate $\hat{f}$, a lower price implies a higher risk-neutral default probability and therefore a higher risk-neutral probability of good assets defaulting. Replacing (44), (39) and (40) into (38):

$$r = \frac{\eta}{\alpha}\left((\alpha-1)\left(1 + \frac{\hat{\phi}\left(1-\hat{f}\right)}{(1-\pi)\left(p^*\alpha - \hat{\phi}\right)}\right) + 1 - \frac{1-\hat{f}}{1-\pi}\right) \tag{45}$$

[10] See Appendix C for details.

[11] Note that if $p^* = \frac{\hat{f}\hat{\phi}+\left(1-\hat{f}\right)}{\alpha}$ then equation (44) implies $\pi = \pi^R$.

As before, π is unobserved, but by taking the derivative of (45) it is possible to establish whether setting $\pi = 0$ gives an upper or a lower bound on r:

$$\frac{\partial r}{\partial \pi} = \frac{\frac{\eta}{\alpha}\left(1-\hat{f}\right)}{(1-\pi)^2}\left[\frac{(\alpha-1)\,\hat{\phi}}{p^*\alpha-\hat{\phi}}-1\right]$$

so if

$$\Upsilon \equiv \frac{(\alpha-1)\,\hat{\phi}}{p^*\alpha-\hat{\phi}}-1>0$$

then setting $\pi = 0$ gives a lower bound, and vice versa. In my sample, the weighted average price of bonds is $p^* = 0.671$, which gives $\Upsilon = -0.939$, implying that $\pi = 0$ gives an upper bound. Replacing $\pi = 0$ in formula (45) gives an upper bound of $r = 0.214$.

5.4 Non-Binary Payoffs

I now assume that the assets' payoffs $q(i)$ take a continuum of possible values instead of just two. However, I maintain the assumption that all good assets have the same payoff (i.e. $q(i) = 1$ for all $i \geq \lambda$); only bad assets' payoffs take a continuum of possible values. This makes sense in the case of debt securities like junk bonds: they have a well-defined upper bound (which is to pay back principal and interest in full) but recovery values upon default are heterogeneous.

I maintain for now the assumption that households know the payoff exactly.[12] In order to describe banks' information, I let assets lie on a two-dimensional continuum $(i,j) \in [0,1]\times[0,1]$ (though payoffs depend only on the i dimension) and let bank θ observe:

$$x(i,j,\theta) = \begin{cases} 1 & \text{if } i \geq \lambda \text{ or } j \geq \theta \\ 0 & \text{otherwise} \end{cases} \tag{46}$$

As in the baseline case, more expert banks make fewer mistakes. Formula (46) implies that when banks do mistake a bad asset for a good one, they select on the payoff-irrelevant dimension j instead of the payoff-relevant dimension i. In other words, bad assets with a higher recovery value are just as likely to be mistaken for good assets as those with a lower recovery value. It is hard to assess whether this assumption is reasonable or not. Quantitatively, it does not seem to make much difference relative to assuming that bad

[12] If households only knew whether $i \geq \lambda$ but not the exact i, this is equivalent to the model with binary payoffs.

assets with higher recovery values are more easily mistaken for good assets.

I will characterize a single-price equilibrium where all good assets that are offered on sale are indeed sold (i.e. are not rationed) and then find conditions such that this is in fact an equilibrium. As in the baseline model, a household of type s who holds asset i will be willing to sell it at price p if $\beta(s)\,q(i) < p$. Therefore, the total supply of i-assets will be

$$s^*(i,p) = \beta^{-1}\left(\frac{p}{q(i)}\right) \tag{47}$$

Since bank θ accepts a fraction $1-\theta$ of bad assets and all good assets, the average quality it buys is:

$$\frac{(1-\theta)\int_0^\lambda s^*(i,p)\,q(i)\,di + \int_\lambda^1 s^*(i,p)\,q(i)\,di}{(1-\theta)\int_0^\lambda s^*(i,p)\,di + \int_\lambda^1 s^*(i,p)\,di}$$

and the indifference condition for the marginal bank is:

$$p^* = \frac{(1-\theta^*)\int_0^\lambda s^*(i,p^*)\,q(i)\,di + \int_\lambda^1 s^*(i,p^*)\,q(i)\,di}{(1-\theta^*)\int_0^\lambda s^*(i,p^*)\,di + \int_\lambda^1 s^*(i,p^*)\,di} \tag{48}$$

Bank θ buys:

$$\frac{\int_\lambda^1 s^*(i,p)\,q(i)\,di}{(1-\theta)\int_0^\lambda s^*(i,p)\,di + \int_\lambda^1 s^*(i,p)\,di}$$

good assets per unit of wealth, so the market-clearing-for-good-assets condition is:

$$p^* = \int_{\theta^*}^{1}\left(\frac{1}{(1-\theta)\int_0^\lambda s^*(i,p^*)\,di + \int_\lambda^1 s^*(i,p^*)\,di}\right)dW(\theta) \tag{49}$$

In order to check that (p^*,θ^*) given by the solution to (48) and (49) does constitute a single-price equilibrium, one must verify that no bank has an incentive to buy assets at any lower price. Since bad assets are partially rationed in market p^* but good assets are not, the supply of assets at any $p < p^*$ will consist of all the bad assets that households are willing to sell at price p, and no good assets. The average quality of this pool will be

$$\bar{q}(p) = \frac{\int_0^\lambda s^*(i,p)\,q(i)\,di}{\int_0^\lambda s^*(i,p)\,di}$$

Banks of any expertise will all obtain an average quality $\bar{q}(p)$ if they buy at price p because, with information given by (46), expertise does not help distinguish among bad assets. Therefore there is a single-price equilibrium if $\bar{q}(p) < p$ for all $p < p^*$, i.e. if there is Akerlof-style unraveling once the good assets are removed. This can be verified numerically for any given example.[13]

In Appendix B I show how to compute the marginal social and private value in this extended model, and hence how to compute r. Importantly, Proposition 1 extends to this case, so one can meaningfully refer to r as a single number. I also show how to compute α, f and η. It is no longer the case that r reduces to a simple function of α, f and η. Therefore measuring r requires quantifying the entire model. I do this as follows.

Mora (2015) reports the distribution of recovery values for bonds that default. I set $q(i)$ for $i < \lambda$ to match this empirical distribution.[14] The resulting function is shown in Figure 5. The function is not steep near $i = 0$, indicating a large quantity of bonds with near-zero recovery values. Approximately 20% of bonds that default have recovery values less than 0.1.

[13] In the binary-payoff case, $\bar{q}(p) = \phi V$ for all $p < p^*$ so this condition does not hold, and there is a two-price equilibrium.

[14] If $F(q)$ is the cdf of recovery values conditional on default, then

$$q(i) = \begin{cases} F^{-1}\left(\frac{i}{\lambda}\right) & \text{if } i < \lambda \\ 1 & \text{if } i \geq \lambda \end{cases}$$

I approximate F with a piecewise linear function on 20 equally-spaced intervals between 0 and 1.5. I thank Nada Mora for kindly sharing her data with me.

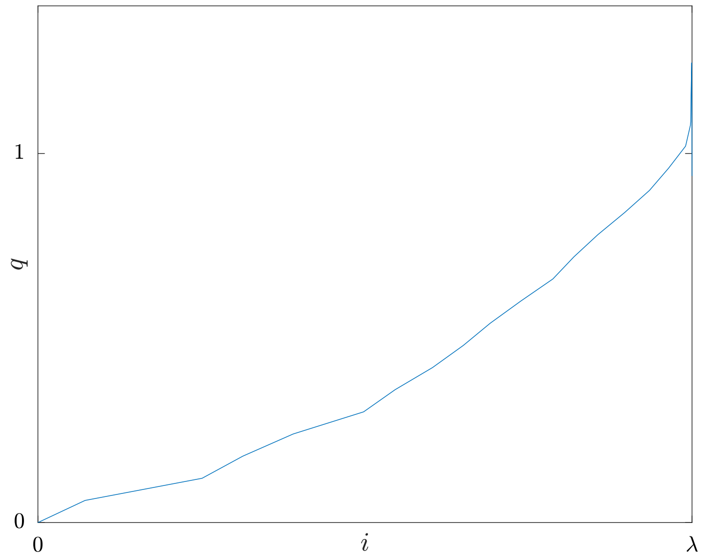

Figure 5: $q(i)$ function that matches empirical distribution of recovery values.

The fraction of bad assets λ and the wealth distribution $W(\theta)$ do not matter separately: an economy with a lower proportion of bad assets is exactly isomorphic to one where the wealth distribution is shifted to the right, so that all banks can reject some bad assets. For this reason, I set λ at the arbitrary value of $\lambda = 0.5$. I assume that $\beta(s)$ and $w(\theta)$ have the following functional forms:

$$\beta(s) = \beta_{\max} - a_1(1-s)$$
$$w(\theta) = a_2 + a_3(1-\theta)$$

This leaves four parameters. I consider different possible values of $\beta_{\max}$, including $\beta_{\max} > 1$, which means that some trades may destroy value. For each $\beta_{\max}$, I set the other three parameters so that the model matches the measured values of α, f and η. Other things being equal, a higher value of a_1 means that βs are more heterogeneous. This in turn implies that there are few marginal sellers, which means lower elasticity η; a high value of a_2 means more high expertise buyers, which drives up prices, drawing in marginal sellers and driving out marginal buyers, which improves the pool of assets sold, lowering the failure rate f; a high value of a_3 means that there is relatively more wealth close to the marginal buyer that at the top of the expertise range, which leads to lower average profitability α.

For each combination of parameters, I compute r in the full model and then compare it to the value that results from direct application of formula (1), which is always the same by construction. In all cases, I verify that $\bar{q}(p) < p$ for all $p < p^*$ so the single-price equilibrium is indeed an equilibrium.

Figure 6 reports the values of r that result from different choices for $\beta_{\max}$. The main consequence of allowing for continuous $q(i)$ is that marginal changes in expertise affect the fraction of bad assets that end up trading (because more expert banks reject them) and, unlike in the baseline model, this matters for social surplus. If $\beta_{\max} < 1$, then these trades of bad assets always create social surplus, so preventing them from happening is undesirable. As a result, r is lower than what formula (1) suggests. Instead, if $\beta_{\max} > 1$, some trades of bad assets are surplus-destroying, so preventing them is an additional social value of marginal increases in expertise, and r can be higher than what formula (1) suggests. Still, for a fairly wide range of values of $\beta_{\max}$, the values of r from the full continuous model are relatively close to formula (1), so the the formula is not too bad of an approximation to what happens in the full model.

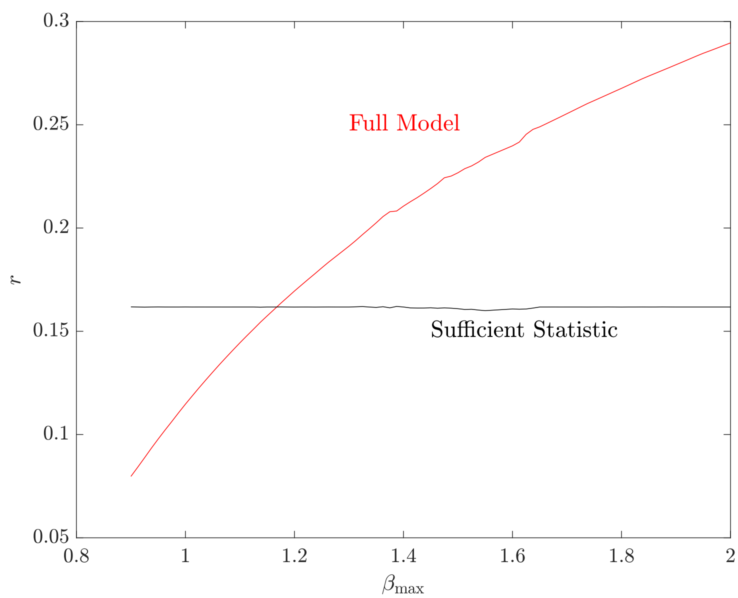

Figure 6: r in the continuous $q(i)$ case

5.5 Failures of Good Assets with Non-Binary Payoffs

I now relax both the assumption of binary payoffs and the assumption of perfectly informed households simultaneously. I assume that there is a probability π that assets that households think are good will fail, in which case their payoffs are a random draw from the same distribution as known-to-be-bad assets. This implies that the expected payoff of an asset $i \geq \lambda$ is $q(i) = \pi\bar{q} + (1-\pi)$. Other than for this change, the equilibrium of the model is found in the same way as in the model with no defaults from good assets. Equations (48) and (49) still determine equilibrium and the formulas for r, α, f and η in Appendix B are still valid.

I fix $\beta_{\max} = 1$ and consider a range of values of π. As in Section 5.4, for each value of π is choose values of a_1, a_2 and a_3 so that the model matches the measured values of α, f and η, and compute r. Figure 7 reports the values of r that result from different values of π. As in the case with binary payoffs, allowing for the possibility that good assets default lowers r, strengthening the conclusion that $r < 1$.

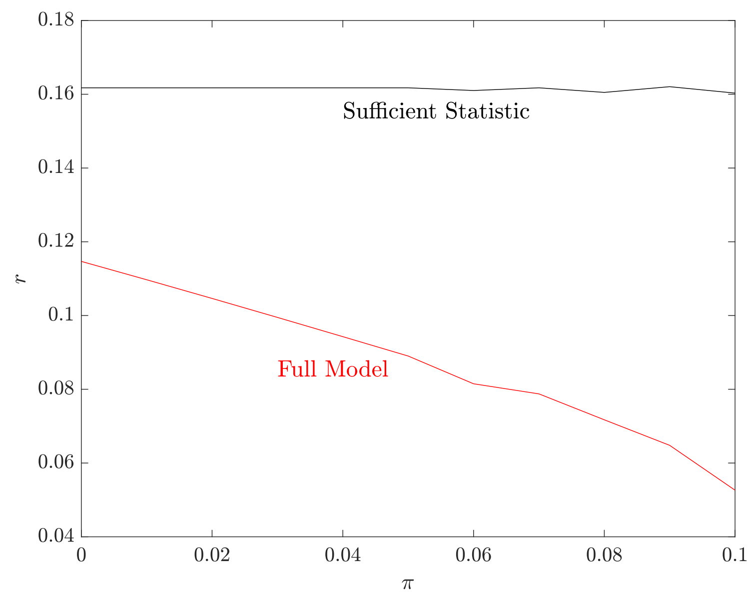

Figure 7: r in the continuous $q(i)$ case with defaults from good assets

6 Discussion

The method I use to measure r has both advantages and limitations, some of which have to do with the method itself and others with the application.

One advantage is that it does not require estimating or making assumptions about the nature of the cost function $c_j(\theta)$ ("how many physicists with PhDs does it take to value a mortgage-backed security?"). Simply assuming that θ is chosen optimally makes it possible to sidestep this question. Another advantage, common to methods based on sufficient statistics, is that the ingredients of r can be measured without measuring all the structural parameters of the model. Chetty (2008) offers a discussion of this type of approach.

One disadvantage, also common to sufficient statistics methods, is that r is a purely local measure at the equilibrium. If some policy were to result in a different equilibrium, then r at the new equilibrium might be different. If one wanted calculate the optimal rate of a simple Pigouvian tax to align private and social incentives it would be necessary to know r at the new equilibrium rather than at the original equilibrium.

Another limitation is that r measures the size of the wedge between $S_j'(\theta_j)$ and $w_j\tau'(\theta_j)$ but not the distance between the equilibrium θ_j and the social optimum θ^{opt} in Figure 2. In order to assess this, it would be necessary to know more about the cost function. For instance, if the marginal cost of expertise increased very steeply, then even a large wedge between r and 1 would imply a small difference between θ_j and θ^{opt}.

The application to junk bonds is predicated on the assumption that this is indeed an asymmetric information environment and that underwriters' role is to mitigate investors' informational disadvantage. The empirical findings on the elasticity of prices and default rates to capital inflows are consistent with this view, but are not definitive proof. Ultimately, all statements about r are conditional on the soundness of the model.

In interpreting the measured values of r, it's important to bear in mind that evaluating trades in environments with asymmetric information is just one of the many things that financial firms do. Therefore the measured r is informative about the net social value of dedicating resources to these types of activities within finance and not necessarily about the industry as a whole. Indeed, the method for measuring r could be applied to businesses that are not usually classified as finance but also involve expertise for trading under asymmetric information, such as used car dealerships.

A maintained assumption is that $(1-\beta(s))V$ represents the social value of the gains from trade. If the trade itself generates externalities then the social gains from trade should be adjusted accordingly. A firm that finances a buyout by issuing junk bonds could bring about new management techniques that other firms learn from or could be destroying value to take advantage of tax benefits. Taking this into account could make the social value of financial expertise higher or lower than measured.

References

Akerlof, G. A.: 1970, The market for "lemons": Quality uncertainty and the market mechanism, *The Quarterly Journal of Economics* **84**(3), 488–500.

Altman, E. I.: 1992, Revisiting the high-yield bond market, *Financial Management* pp. 78–92.

Altman, E. I., Brady, B., Resti, A. and Sironi, A.: 2005, The link between default and recovery rates: Theory, empirical evidence, and implications, *The Journal of Business* **78**(6), 2203.

Arnold, L. G. and Riley, J. G.: 2009, On the possibility of credit rationing in the stiglitz-weiss model, *American Economic Review* **99**(5), 2012–21.

Bolton, P., Santos, T. and Scheinkman, J. A.: 2011, Cream skimming in financial markets. NBER working paper No. 16804.

Booth, J. R. and Smith, II, R. L.: 1986, Capital raising, underwriting and the certification hypothesis, *Journal of Financial Economics* **15**(1-2), 261 – 281.

Brewer, E. and Jackson, W. E.: 2000, Requiem for a market maker: the case of drexel burnham lambert and junk bonds, *Journal of Financial Services Research* **17**(3), 209–235.

Chetty, R.: 2008, Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. NBER working paper No. 14399.

Cochrane, J. H.: 2013, Finance: Function Matters, Not Size, *Journal of Economic Perspectives* **27**(2), 29–50.

Datta, S., Iskandar-Datta, M. and Patel, A.: 1997, The pricing of initial public offers of corporate straight debt, *the Journal of Finance* **52**(1), 379–396.

Fishman, M. J. and Parker, J. A.: 2015, Valuation, adverse selection, and market collapses, *Review of Financial Studies* .

Gale, D.: 1996, Equilibria and pareto optima of markets with adverse selection, *Economic Theory* **7**(2), 207–235.

Gande, A., Puri, M. and Saunders, A.: 1999, Bank entry, competition, and the market for corporate securities underwriting, *Journal of Financial Economics* **54**(2), 165–195.

Giesecke, K., Longstaff, F. A., Schaefer, S. and Strebulaev, I.: 2011, Corporate bond default risk: A 150-year perspective, *Journal of Financial Economics* pp. 233–250.

Glode, V., Green, R. C. and Lowery, R.: 2012, Financial expertise as an arms race, *The Journal of Finance* **67**(5), 1723–1759.

Greenwood, R. and Scharfstein, D.: 2013, The growth of finance, *The Journal of Economic Perspectives* pp. 3–28.

Guerrieri, V., Shimer, R. and Wright, R.: 2010, Adverse selection in competitive search equilibrium, *Econometrica* **78**(6), 1823–1862.

Jewell, J. and Livingston, M.: 1998, Split ratings, bond yields, and underwriter spreads, *Journal of Financial Research* **21**(2), 185–204.

Kurlat, P.: 2016, Asset markets with heterogeneous information, *Econometrica* **84**, 33–85.

Levine, R.: 1997, Financial development and economic growth: Views and agenda, *Journal of Economic Literature* **35**(2), 688–726.

Levine, R.: 2005, Finance and growth: theory and evidence, *Handbook of economic growth* **1**, 865–934.

Mora, N.: 2015, Creditor recovery: The macroeconomic dependence of industry equilibrium, *Journal of Financial Stability* **18**, 172–186.

Murphy, K. M., Shleifer, A. and Vishny, R. W.: 1991, The allocation of talent: Implications for growth, *The Quarterly Journal of Economics* **106**(2), 503–30.

Philippon, T.: 2010, Financiers versus engineers: Should the financial sector be taxed or subsidized?, *American Economic Journal: Macroeconomics* **2**(3), 158–82.

Philippon, T.: 2014, Has the us finance industry become less efficient? on the theory and measurement of financial intermediation. NYU working paper.

Philippon, T. and Reshef, A.: 2012, Wages and human capital in the u.s. finance industry: 1909-2006, *The Quarterly Journal of Economics* .

Reilly, F. K., Wright, D. J. and Gentry, J. A.: 2009, Historic changes in the high yield bond market, *Journal of Applied Corporate Finance* **21**(3), 65 – 79.

Shakhnov, K.: 2014, The allocation of talent: Finance versus entrepreneurship. EUI working paper.

Stewart, J. B.: 1992, *Den of thieves*, Simon and Schuster.

Stiglitz, J. E. and Weiss, A.: 1981, Credit rationing in markets with imperfect information, *American Economic Review* **71**(3), 393–410.

Tobin, J.: 1984, On the efficiency of the financial system, *Lloyds Bank Review* .

Wilson, C.: 1980, The nature of equilibrium in markets with adverse selection, *The Bell Journal of Economics* **11**(1), 108–130.

A Equilibrium in the Baseline Model

A.1 Equilibrium Definition

Each possible price $p \in [0, V]$ defines a market and any asset can in principle be traded in any market. Markets need not clear: assets that are offered for sale in market p may remain totally or partially unsold.

Households trade by choosing at what prices to put their asset on sale. Markets are non-exclusive: households are allowed to offer their asset for sale at as many prices as they want. This implies that a household of type s who owns asset i will simply choose a reservation price $p^R(i, s)$ and put its asset on sale at every $p \geq p^R(i, s)$ and not at any price below that.[15] From the household's point of view, the only thing that matters about the equilibrium is at what price it's possible to sell its asset i, i.e. the extent to which it will face rationing at each price. Formally, this is captured by a "rationing function" $\mu : [0, V] \times [0, 1] \to \mathbb{R}$. $\mu(p, i)$ is the number of assets that a household would end up selling if it offers one unit of asset i on sale with the reservation price p (thereby offering it on sale at every price in $[p, V]$). Implicit in this formulation is the assumption that assets are perfectly divisible, so there is exact pro-rata rationing rather than a probability of selling an indivisible unit.

[15]There is an extra assumption involved in this. There will be many prices at which it's impossible to sell assets so the household is indifferent between offering its asset on sale in them or not. A reservation price is the only optimal strategy that is robust to a small chance of selling at every price.

A household of type s who owns asset i solves:

$$\max_{p^R} \quad \int_{p^R}^{V} p d\mu\left(p,i\right) + \left[1-\mu\left(p^R,i\right)\right]\beta\left(s\right)q\left(i\right) \tag{50}$$

$$s.t. \quad \mu\left(p^R,i\right) \leq 1 \tag{51}$$

The first term in (50) represents the proceeds from selling the asset, possibly fractionally and across many prices. The second term represents the dividends obtained from whatever fraction of the asset the household retains. Constraint (51) limits the household to not sell more than one unit in total.

This problem has a simple solution. Define

$$p^L\left(i\right) \equiv \max\left\{\inf\left\{p:\mu\left(p,i\right)<1\right\},0\right\}$$

$p^L\left(i\right)$ is the highest reservation price that a household can set and still be sure to sell its entire asset; if there is no positive price that guarantees selling the entire asset, then $p^L\left(i\right)=0$. It's immediate that the solution to program (50) is:

$$p^R\left(i,s\right) = \max\left\{p^L\left(i\right),\beta\left(s\right)V\right\} \tag{52}$$

If it's possible to sell the entire asset at a price above the household's own valuation, then the household sets the reservation price at the level that guarantees selling; otherwise the reservation price is the household's own valuation.

Turn now to the bank's problem. It has two stages: first the bank chooses a level of expertise and then it trades assets. In the second stage, the bank trades by choosing a quantity δ, a price p and an acceptance rule χ. An acceptance rule is a function $\chi:[0,1]\to\{0,1\}$ from the set of assets to $\{0,1\}$, where $\chi\left(i\right)=1$ means that the bank is willing to accept asset i and $\chi\left(i\right)=0$ means it is not. By trading in market p with acceptance rule χ, the bank obtains χ-acceptable assets in proportion to the quantities that are offered on sale at price p. A bank may only impose acceptance rules that are informationally feasible given the expertise it has acquired, so it cannot discriminate between assets that it cannot tell apart, i.e. $\chi\left(i\right)=\chi\left(i'\right)$ whenever $x\left(i,\theta\right)=x\left(i',\theta\right)$.

From the point of view of banks, the only thing that matters about the equilibrium is what distribution of assets it will obtain for each possible combination of price and acceptance rule it could choose. Formally, this is captured by a measure $A\left(\cdot;\chi,p\right)$ on the set of assets $[0,1]$ for each χ,p. For any subset $I\subseteq[0,1]$, $A\left(I;\chi,p\right)$ is the measure of assets $i\in I$ that a bank will end up with if it demands one unit at price p with acceptance rule χ.

Therefore in the trading stage, a bank with expertise θ and wealth w solves:

$$\max_{\delta,p,\chi} \quad \delta \left[\int\limits_{[0,1]} q(i)\, dA(i;\chi,p) - pA([0,1];\chi,p) \right] \tag{53}$$

$$s.t. \quad \delta pA([0,1];\chi,p) \leq w \tag{54}$$

$$\chi(i) = \chi(i') \quad \text{whenever} \quad x(i,\theta) = x(i',\theta) \tag{55}$$

(53) adds all the dividends $q(i)$ of the assets the bank buys, subtracts what it pays per unit and multiplies by total demand δ; (54) is the budget constraint and (55) imposes that the bank use an informationally feasible acceptance rule.

Notice that w enters the problem only in the budget constraint, which is linear. This implies that δ will be linear in w and p and χ will not depend on w. Let $\delta(\theta)$, $p(\theta)$ and $\chi(\theta)$ denote the solution to the bank's problem for a bank with $w = 1$ and expertise θ, and let $\tau(\theta)$ be the maximized value of (53) for $w = 1$.

The two key equilibrium objects are the rationing function $\mu(p,i)$ and the allocation measures $A(\cdot;\chi,p)$. The allocation measures $A(\cdot;\chi,p)$ formalize the notion that banks obtain representative samples from the assets on sale that they find acceptable. The rationing function μ formalizes the notion that whether assets that are put on sale are actually sold depends on how many units are demanded by banks who find them acceptable. To compute A and μ, first define supply and demand.

The supply of asset i at price p is:

$$S(i;p) = \int\limits_{s} \mathbb{I}\left(p^R(i,s) \leq p\right) \tag{56}$$

(56) is just aggregating all the supply from households whose reservation prices are below p.

Demand is defined as a measure. Suppose X is some set of possible acceptance rules. Define

$$\Theta(X,p) \equiv \{\theta : \chi(\theta) \in X, p(\theta) \geq p\}$$

$\Theta(X,p)$ is the set of bank types who choose to buy at prices above p using acceptance rules in the set X. Aggregating $\delta(\theta)$ over this set gives demand:

$$D(X,p) = \int\limits_{\theta \in \Theta(X,p)} \delta(\theta)\, dW(\theta) \tag{57}$$

One complication is that if different banks impose different acceptance rules in the same market, the allocation will depend on the order in which they execute their trades because each successive bank will alter the sample from which the following banks draw assets. Kurlat (2016) shows that if one allows markets for each of the possible orderings and lets traders self-select, then in equilibrium trades will take place in a market where the less restrictive banks execute their trades first.[16] Less-restrictive banks' trades do not alter the relative proportions of acceptable assets available for the more-restrictive banks who follow them so, as long as acceptable assets don't run out, all bankers obtain assets as though they were drawing from the original sample. This means that (as long as acceptable assets don't run out before a bank with rule acceptance rule χ trades, which does not happen in equilibrium) the density of measure $A(\cdot;\chi,p)$ is:

$$a(i;\chi,p) = \begin{cases} \frac{\chi(i)S(i;p)}{\int \chi(i)S(i;p)di} & \text{if } \int \chi(i)\, S(i;p)\, di > 0 \\ 0 & \text{otherwise} \end{cases} \tag{58}$$

Knowing A, the rationing faced by an asset i depends on the the ratio of the total demand that gets satisfied (added across all χ) to supply, so

$$\mu(p,i) = \int\limits_{\substack{\tilde{p} \geq p \\ \text{all } \chi}} \frac{a(i;\chi,\tilde{p})}{S(i;\tilde{p})} dD(\chi,\tilde{p}) \tag{59}$$

I define equilibrium in two steps. First I define a conditional equilibrium, i.e. an equilibrium given the first-stage choices by banks that result in $W(\theta)$.

Definition 1. Taking $W(\theta)$ as given, a conditional equilibrium is given by reservation prices $p^R(i,s)$, buying plans $\{\delta(\theta), p(\theta), \chi(\theta)\}$, rationing measures $\mu(\cdot;i)$ and allocation measures $A(\cdot;\chi,p)$ such that: $p^R(i,s)$ solves the household's problem for all i, s, taking $\mu(\cdot,i)$ as given;

[16] An acceptance rule $\tilde{\chi}$ is less restrictive than another rule χ if $\chi(i) = 1$ implies $\tilde{\chi}(i) = 1$ but there exists some i such that $\tilde{\chi}(i) = 1$ and $\chi(i) = 0$. Under the information structure (3), all feasible acceptance rules can be ranked by restrictiveness.

$\{\delta(\theta), p(\theta), \chi(\theta)\}$ solves the bank's second stage problem for all θ, taking $A(\cdot;\chi,p)$ as given and $\mu(\cdot;i)$ and $A(\cdot;\chi,p)$ satisfy the consistency conditions (58) and (59).

Using this, I now define a full equilibrium. The usefulness of this two-step definition is that it is possible to focus on characterizing the conditional equilibrium without fully specifying the cost functions c_j that govern the banks' first-stage decisions.

Definition 2. An equilibrium is given by expertise choices θ_j, a wealth distribution $W(\theta)$ and a conditional equilibrium $\left\{p^R, \delta, p, \chi, \mu, A\right\}$ such that: θ_j solves the bank's first stage problem for all j, taking the conditional equilibrium as given; $W(\theta)$ is defined by (5) and $\left\{p^R, \delta, p, \chi, \mu, A\right\}$ is a conditional equilibrium given $W(\theta)$.

A.2 Equilibrium Characterization

Taking $W(\theta)$ as given, let p^*, θ^* and s^* be the highest-p^* solution to the system of equations (6),(8) and (9). Furthermore, assume the following:

Assumption 1. $\frac{1}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)}{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta^*)}V < 1$ for all $p > p^*$

Proposition 2. *If Assumption 1 holds, there is a unique conditional equilibrium, where:*

1. Reservation prices are:

$$p^R(i,s) = \begin{cases} \max\{p^*, \beta(s)V\} & \text{if } i \geq \lambda \\ 0 & \text{if } i < \lambda \end{cases} \tag{60}$$

2. The solution to the banks' problem is:

$$\{\delta(\theta), p(\theta), \chi(\theta)\} = \begin{cases} \left\{\frac{1}{p^*}, p^*, \mathbb{I}(i \geq \lambda\theta)\right\} & \text{if } \theta \geq \theta^* \\ \{0,0,0\} & \text{if } \theta < \theta^* \end{cases} \tag{61}$$

3. The allocation function is:

$$a(i;\chi,p) = \begin{cases} \frac{\beta^{-1}\left(\frac{p}{V}\right)\chi(i)}{\int_0^\lambda \chi(i)di + \int_\lambda^1 \chi(i)\beta^{-1}\left(\frac{p}{V}\right)di} & \text{if } i \geq \lambda \text{ and } p \geq p^* \\ \frac{\chi(i)}{\int_0^\lambda \chi(i)di + \int_\lambda^1 \chi(i)\beta^{-1}\left(\frac{p}{V}\right)di} & \text{if } i < \lambda \text{ and } p \geq p^* \\ 0 & \text{if } i \geq \lambda \text{ and } p < p^* \\ \frac{\chi(i)}{\int_0^\lambda \chi(i)di} & \text{if } i < \lambda \text{ and } p < p^* \end{cases} \tag{62}$$

4. The rationing function is:

$$\mu(p,i)=\begin{cases}1 & \text{if } i\geq\lambda, p\leq p^{*}\\ \int_{\theta^{*}}^{\frac{i}{\lambda}}\frac{1}{\lambda(1-\theta)+s^{*}(1-\lambda)}\frac{1}{p^{*}}dW(\theta) & \text{if } i\in[\lambda\theta,\lambda), p\leq p^{*}\\ 0 & \text{if } i<\lambda\theta, p\leq p^{*}\\ 0 & \text{if } p>p^{*}\end{cases} \tag{63}$$

Proof.

(a) **Equations** (60)-(63) **constitute an equilibrium.**

i. Household optimization. (63) implies that:

$$p^{L}(i)=\begin{cases}p^{*} & \text{if } i\geq\lambda\\ 0 & \text{if } i<\lambda\end{cases}$$

This immediately implies that $p^{R}(i,s)$ from (60) solves the household's problem.

ii. Bank optimization.

A. $\chi(\theta)$ is the optimal acceptance rule because, given (62), any other rule that satisfies (55) includes a higher proportion of bad assets.

B. At any $p<p^{*}$, there are no good assets on sale so it is not optimal for any bank to choose this. For any $p>p^{*}$:

$$\begin{aligned}\frac{1}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta^{*})}&<\frac{1}{p^{*}}\frac{s^{*}}{s^{*}(1-\lambda)+\lambda(1-\theta^{*})}\\ \frac{p^{*}}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{s^{*}}&<\frac{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta^{*})}{s^{*}(1-\lambda)+\lambda(1-\theta^{*})}\\ \frac{p^{*}}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{s^{*}}&<\frac{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta)}{s^{*}(1-\lambda)+\lambda(1-\theta)}\quad\text{for all }\theta\geq\theta^{*}\\ \frac{1}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta)}&<\frac{1}{p^{*}}\frac{s^{*}}{s^{*}(1-\lambda)+\lambda(1-\theta)}\quad\text{for all }\theta\geq\theta^{*}\end{aligned} \tag{64}$$

The first step is Assumption (1); the second is just rearranging; the third follows because the right hand side is increasing in θ and the last is just rearranging. Inequality (64) implies that all banks with $\theta\geq\theta^{*}$ prefer to buy at price p^* than at higher prices. Therefore if they buy at all they buy at price p^*.

C. For $\theta > \theta^*$, $\tau(\theta) > 0$ so the budget constraint (54) binds; for $\theta < \theta^*$ there is no $\chi(\theta)$ that satisfies (55) and leads to a positive value for the objective (53). Therefore $\delta(\theta)$ is optimal .

iii. Consistency of A and μ. Replacing reservation prices (60) into (56) and using this to replace $S(i;p)$ into (58) leads to (62). Adding up demand using (61) and (57) and replacing in (59) implies (63).

(b) **The equilibrium is unique**

Note first that since no feasible acceptance rule has $\chi(i) \neq \chi(i')$ for $i, i' \geq \lambda$, this implies that $p^L(i) = p^L(\lambda)$ and $S(i,p) = S(\lambda, p)$ for all $i \geq \lambda$. Now proceed by contradiction.

Suppose there is another equilibrium with $p^L(\lambda) < p^*$. Households' optimization condition (52) and formula (56) for supply imply that for $p \in \left[p^L(\lambda), p^*\right]$:

$$S(i,p) = \begin{cases} \beta^{-1}\left(\frac{p}{V}\right) & \text{if } i \geq \lambda \\ 1 & \text{if } i < \lambda \end{cases} \tag{65}$$

(65) implies that all banks with $\theta > \theta^*$ can attain $\tau(\theta) > 0$ by choosing p^*. By (64), they prefer p^* to any $p' > p^*$ and therefore in equilibrium they all chose some $p(\theta) \in \left[p^L(\theta), p^*\right]$ and $\delta(\theta) = \frac{1}{p(\theta)}$. Using (58):

$$a(i, \chi(\theta), p(\theta)) = \frac{\beta^{-1}\left(\frac{p(\theta)}{V}\right)}{\beta^{-1}\left(\frac{p(\theta)}{V}\right) + \lambda(1-\theta)} \quad \text{for all } i \geq \lambda$$

Using (59), this implies that

$$\mu(p, \lambda) = \int_{\{\theta: p(\theta) \geq p\}} \frac{1}{\beta^{-1}\left(\frac{p(\theta)}{V}\right) + \lambda(1-\theta)} \frac{1}{p(\theta)} dW(\theta)$$

and therefore

$$\mu\left(p^L(\lambda),\lambda\right) \geq \int_{\theta^*}^{1} \frac{1}{\beta^{-1}\left(\frac{p(\theta)}{V}\right)+\lambda(1-\theta)}\frac{1}{p(\theta)}dW(\theta)$$
$$\geq \int_{\theta^*}^{1} \frac{1}{s^*+\lambda(1-\theta)}\frac{1}{p^*}dW(\theta)$$
$$= 1 \tag{66}$$

The first inequality follows because the set $\{\theta : p(\theta) \geq p(\lambda)\}$ includes $[\theta^*, 1]$; the second follows because $\beta^{-1}\left(\frac{p^*}{V}\right) = s^*$, β^{-1} is increasing and $p^* \geq p(\theta)$; the last equality is just the market clearing condition (9). Furthermore, if $p(\theta) < p^*$ for a positive measure of banks, then (66) is a strict inequality, which leads to a contradiction. Instead, if $p(\theta) = p^*$ for almost all banks, then $p^L(\lambda) = p^*$, which contradicts the premise.

Suppose instead that there is an equilibrium such that $p^L(\lambda) > p^*$. This implies that there is no supply of good assets at any price $p < p^L(\lambda)$ and therefore no bank with $\theta < \theta^*$ chooses $\delta(\theta) > 0$ and banks $\theta \in [\theta^*, 1]$ choose some price $p(\theta) \geq p^L(\lambda)$ and $\delta(\theta) \leq \frac{1}{p(\theta)}$. Therefore, using (58) and (59), we have

$$\mu\left(p^L(\lambda),\lambda\right) \leq \int_{\theta^*}^{1} \frac{1}{\beta^{-1}\left(\frac{p(\theta)}{V}\right)+\lambda(1-\theta)}\frac{1}{p(\theta)}dW(\theta)$$
$$< \int_{\theta^*}^{1} \frac{1}{s^*+\lambda(1-\theta)}\frac{1}{p^*}dW(\theta)$$
$$= 1$$

The first inequality follows from $\delta(\theta) \leq \frac{1}{p(\theta)}$; the second follows because $\beta^{-1}\left(\frac{p^*}{V}\right) = s^*$, β^{-1} is increasing and $p^* < p(\theta)$; the last equality is just the market clearing condition (9). Again, this is a contradiction.

Therefore any equilibrium must have $p^L(\lambda) = p^*$. The rest of the equilibrium objects follow immediately.

$\square$

A.3 The Role of Assumption 1

The equilibrium concept gives banks the option to buy assets at prices other than p^*. Buying at lower prices is clearly worse than buying at p^* because the reservation price for good assets is at least p^* so no good assets are on sale at lower prices. Assumption 1 ensures that buying at higher price is not preferred either. Given the reservation prices (60), the surplus per unit of wealth for bank θ^* if it buys at price $p > p^*$ is:

$$\frac{1}{p}\left[\frac{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)V}{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta^*)}-p\right]$$

In principle, the bank faces a tradeoff: better selection (because β^{-1} is an increasing function) but a higher price. Assumption 1 ensures that the direct higher-price effect dominates and a bank with expertise θ^* has no incentive to pay higher prices to ensure better selection. It is then possible to show that if this is true for the marginal bank θ^*, it is true for all banks: higher-θ banks care even less about selection because they can filter assets themselves and lower-θ banks can never earn surplus in a market where θ^* would not. One can still solve for equilibria where Assumption 1 does not hold, but they are somewhat more complicated. Wilson (1980), Stiglitz and Weiss (1981) and Arnold and Riley (2009) analyze the implications of models where an analogue of Assumption 1 doesn't hold.

B Continuous $q(i)$

B.1 Computing r

Define:

$$S_L(p) \equiv \int_0^\lambda s^*(i,p)\,di$$

$$S_H(p) \equiv \int_\lambda^1 s^*(i,p)\,di$$

These represent, respectively, the quantity of bad and good assets offered on sale at price p. Further define:

$$Q_L(p) \equiv \int_0^\lambda s^*(i,p)\, q(i)\, di$$

$$Q_H(p) \equiv \int_\lambda^1 s^*(i,p)\, q(i)\, di$$

These represent, respectively, total dividends of bad and good assets offered on sale. Their derivatives are given by:

$$S'_L(p) = \int_0^\lambda \frac{\partial s^*(i,p)}{\partial p} di$$

$$S'_H(p) = \int_\lambda^1 \frac{\partial s^*(i,p)}{\partial p} di$$

$$Q'_L(p) = \int_0^\lambda \frac{\partial s^*(i,p)}{\partial p} q(i)\, di$$

$$Q'_H(p) = \int_\lambda^1 \frac{\partial s^*(i,p)}{\partial p} q(i)\, di$$

The equilibrium conditions (48) and (49)can be rewritten as:

$$p^* = \frac{(1-\theta^*)\, Q_L(p^*) + Q_H(p^*)}{(1-\theta^*)\, S_L(p^*) + S_H(p^*)} \tag{67}$$

$$p^* = \int_{\theta^*}^1 \frac{1}{(1-\theta)\, S_L(p^*) + S_H(p^*)} dW(\theta) \tag{68}$$

and in matrix form:

$$K(p^*, \theta^*) = \begin{pmatrix} p^* - \frac{(1-\theta^*)Q_L(p^*)+Q_H(p^*)}{(1-\theta^*)S_L(p^*)+S_H(p^*)} \\ p^* - \int_{\theta^*}^1 \frac{w(\theta)}{(1-\theta)S_L(p^*)+S_H(p^*)} d\theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The gradient of $K$is:

$$D = \begin{pmatrix} d_{1p} & d_{1\theta} \\ d_{2p} & d_{2\theta} \end{pmatrix}$$

with

$$\begin{aligned}
d_{1p} &= 1 - \frac{\begin{array}{c}\left[\left(1-\theta^*\right) Q_L'\left(p^*\right) + Q_H'\left(p^*\right)\right]\left[\left(1-\theta^*\right) S_L\left(p^*\right) + S_H\left(p^*\right)\right] \\ -\left[\left(1-\theta^*\right) S_L'\left(p^*\right) + S_H'\left(p^*\right)\right]\left[\left(1-\theta^*\right) Q_L\left(p^*\right) + Q_H\left(p^*\right)\right]\end{array}}{\left[\left(1-\theta^*\right) S_L\left(p^*\right) + S_H\left(p^*\right)\right]^2} \\
d_{1\theta} &= \frac{Q_L\left(p^*\right) S_H\left(p^*\right) - S_L\left(p^*\right) Q_H\left(p^*\right)}{\left[\left(1-\theta^*\right) S_L\left(p^*\right) + S_H\left(p^*\right)\right]^2} \\
d_{2p} &= 1 + \int_{\theta^*}^1 w\left(\theta\right)\left[\left(1-\theta\right) S_L\left(p^*\right) + S_H\left(p^*\right)\right]^{-2}\left(\left(1-\theta\right) S_L'\left(p^*\right) + S_H'\left(p^*\right)\right) d\theta \\
d_{2\theta} &= \frac{w\left(\theta^*\right)}{\left(1-\theta^*\right) S_L\left(p^*\right) + S_H\left(p^*\right)}
\end{aligned}$$

In equilibrium, bad assets will be rationed and, in order to compute how much surplus is created (or destroyed) from trades of bad assets, I need to compute the fraction μ of bad assets put on sale that will actually be sold. Buyer θ will buy a total of

$$\frac{1}{p}\frac{\left(1-\theta\right) S_L\left(p\right)}{\left(1-\theta\right) S_L\left(p\right) + S_H\left(p\right)}$$

bad assets per unit of wealth, so total demand of bad assets will be

$$\frac{1}{p}\int_{\theta^*}^{1} \frac{\left(1-\theta\right) S_L\left(p\right)}{\left(1-\theta\right) S_L\left(p\right) + S_H\left(p\right)} dW\left(\theta\right)$$

and therefore

$$\mu = \frac{1}{p^*}\int_{\theta^*}^{1} \frac{\left(1-\theta\right)}{\left(1-\theta\right) S_L\left(p^*\right) + S_H\left(p^*\right)} dW\left(\theta\right) \tag{69}$$

The total surplus will be

$$S = \underbrace{\mu\left(p^*, \theta^*\right)}_{\text{rationing}} \int_0^\lambda \left[\underbrace{q\left(i\right)\int_0^{s\left(i,p^*\right)} \left(1-\beta\left(s\right)\right) ds}_{\text{gains from trade in asset } i<\lambda}\right] di + \int_\lambda^1 \left[\underbrace{q\left(i\right)\int_0^{s\left(i,p^*\right)} \left(1-\beta\left(s\right)\right) ds}_{\text{gains from trade in asset } i\geq\lambda}\right] di$$

Taking the derivative, the marginal social value of an increase in expertise is given by:

$$S'\left(\theta_j\right) = \mu\left(p^*, \theta^*\right) \int\limits_0^1 \left(q\left(i\right)\left(1-\beta\left(s^*\left(i,p^*\right)\right)\right) \frac{ds^*\left(i,p^*\right)}{d\theta_j} + q\left(i\right) \frac{d\mu\left(p^*, \theta^*\right)}{d\theta_j} \int_0^{s\left(i,p^*\right)} \left(1-\beta\left(s\right)\right) ds \right) di \tag{70}$$

I need to compute $\frac{ds^*(i,p^*)}{d\theta_j}$. and $\frac{d\mu(p^*,\theta^*)}{d\theta_j}$ to replace in (70). Rewrite $\frac{ds^*(i,p^*)}{d\theta_j}$ as

$$\frac{ds^*\left(i,p^*\right)}{d\theta_j} = \frac{\partial s^*\left(i,p^*\right)}{\partial p^*} \frac{\partial p^*}{\partial \theta_j} \tag{71}$$

The term $\frac{\partial s^*(i,p^*)}{\partial p^*}$ can be computed directly from (47). Using the implicit function theorem:

$$\begin{pmatrix} \frac{\partial p^*}{\partial \epsilon} \\ \frac{\partial \theta^*}{\partial \epsilon} \end{pmatrix} = -D^{-1} \begin{pmatrix} 0 \\ \frac{\partial K_2^*}{\partial \theta_j} \end{pmatrix} \tag{72}$$

where, using (49):

$$\frac{\partial K_2}{\partial \theta_j} = -w_j \left(\left(1-\theta_j\right) S_L\left(p^*\right) + S_H\left(p^*\right)\right)^{-2} S_L\left(p^*\right) \tag{73}$$

Now rewrite $\frac{d\mu(p^*,\theta^*)}{d\theta_j}$ as

$$\frac{d\mu}{d\theta_j} = \underbrace{\frac{\partial \mu}{\partial \theta_j}}_{\text{direct efffect}} + \underbrace{\frac{\partial \mu\left(i;p^*,\theta^*\right)}{\partial p^*} \frac{\partial p^*}{\partial \theta_j}}_{\text{price effect}} + \underbrace{\frac{\partial \mu\left(i;p^*,\theta^*\right)}{\partial \theta^*} \frac{\partial \theta^*}{\partial \theta_j}}_{\text{marginal seller effect}} \tag{74}$$

Using (69), the direct effect is:

$$\frac{\partial \mu}{\partial \theta_j} = -\frac{1}{p^*} \frac{S_H\left(p^*\right)}{\left[\left(1-\theta_j\right) S_L\left(p^*\right) + S_H\left(p^*\right)\right]^2}$$

and the indirect effects are:

$$\frac{\partial \mu\left(i;p^*,\theta^*\right)}{\partial p^*} = -\left[\begin{array}{c} \frac{1}{\left(p^*\right)^2} \int_{\theta^*}^1 \frac{(1-\theta)}{(1-\theta)S_L(p^*)+S_H(p^*)} dW\left(\theta\right) \\ +\frac{1}{p^*} \int_{\theta^*}^1 \frac{\left((1-\theta)S_L'(p^*)+S_H'(p^*)\right)(1-\theta)}{\left((1-\theta)S_L(p^*)+S_H(p^*)\right)^{-2}} dW\left(\theta\right) \end{array} \right]$$

$$\frac{\partial \mu\left(i;p^*,\theta^*\right)}{\partial \theta^*} = -\frac{1}{p^*} \frac{\left(1-\theta^*\right) w\left(\theta^*\right)}{\left(1-\theta^*\right) S_L\left(p^*\right) + S_H\left(p^*\right)}$$

with $\frac{\partial p^*}{\partial \theta_j}$ and $\frac{\partial \theta^*}{\partial \theta_j}$ given by (72) and (73). Replacing (71) and (74) into (70) gives the marginal social surplus.

Profits for bank j are:

$$w_j \tau(\theta_j) = \frac{w_j}{p^*}\left[\frac{(1-\theta) Q_L(p^*) + Q_H(p^*)}{(1-\theta) S_L(p^*) + S_H(p^*)} - p^*\right]$$

so the marginal private gain from increasing expertise is:

$$w_j \tau'(\theta_j) = \frac{1}{p^*}\frac{S_L(p^*) Q_H(p^*) - Q_L(p^*) S_H(p^*)}{\left[(1-\theta_j) S_L(p^*) + S_H(p^*)\right]^2} \tag{75}$$

Taking the ratio of (70) and (75) and simplifying:

$$r = \frac{\left(\begin{array}{c} D_{12}^{-1} S_L(p^*)\,\mu\left[\int_0^\lambda q(i)\left((1-\beta(s^*(i,p^*)))\frac{\partial s^*(i,p^*)}{\partial p^*}\right)di + \int_\lambda^1 q(i)(1-\beta(s^*(i,p^*)))\frac{\partial s^*(i,p^*)}{\partial p^*}di\right] \\ +\left(-\frac{1}{p^*}S_H(p^*) + \left(\frac{\partial \mu(i;p^*,\theta^*)}{\partial p^*}D_{12}^{-1} + \frac{\partial \mu(i;p^*,\theta^*)}{\partial \theta^*}D_{22}^{-1}\right)S_L(p^*)\right)\int_0^\lambda q(i)\left(\int_0^{s(i,p^*)}(1-\beta(s))\,ds\right)di \end{array}\right)}{\frac{1}{p^*}\left(S_L(p^*) Q_H(p^*) - Q_L(p^*) S_H(p^*)\right)}$$

which does not depend on θ_j or w_j, so Proposition 1 holds.

B.2 Computing α, f and η

Banks' average profitability is given by:

$$\alpha = \frac{\mu Q_L + Q_H}{\int_{\theta^*}^1 w(\theta)\,d\theta}$$

The numerator is the total dividends from assets that are actually sold; the denominator is the total funds that buyers spend on assets.

The fraction of bad assets among traded assets is

$$f = 1 - \frac{p^* S_H}{\int_{\theta^*}^1 dW(\theta)}$$

The total number of good assets traded is

$$G = S_H(p^*)$$

so its elasticity with respect to a capital inflow is given by:

$$\eta = \frac{S_H'(p^*)\frac{dp^*}{d\Delta}}{S_H(p^*)}$$

where

$$\begin{aligned}\frac{dp^*}{d\Delta} &= -D_{12}^{-1}\frac{\partial K_2^*}{\partial \Delta} \\ &= D_{12}^{-1}p^*\end{aligned}$$

so

$$\eta = \frac{S_H'(p^*)}{S_H(p^*)}D_{12}^{-1}p^*$$

C Data Sources and Variable Definitions

The Thomson Reuters/Securities Data Company contains data on all corporate bonds issued in the United States. For each bond, the database reports: date of issuance, dollar volume, maturity, coupon rate, yield and price at issuance. From this database I extract all bonds flagged as "high-yield", where high yield is defined as "having a Standard & Poor's rating of BB+ and below or a Moody's rating of Ba1 and below". The Bloomberg database in principle contains the same universe of bonds, and reports the same variables on them. I select from it all the bonds rated below investment grade by either S&P, Moody's or Fitch. Both databases contain bonds that are not included in the other, and in some instances they report inconsistent information about the same bond. I simply add the two databases together, eliminating duplicates and following SDC when there are discrepancies. The date ranges from 1977 to 2010 (by date of issuance). This leaves a total of 30,193 bonds in my main sample, of which I have price information for 17,872.

For each bond, I also record the yield on a Treasury bond of the same maturity at the date of issuance. To construct the yield, I obtain from Bloomberg the Treasury rates of standard maturities (1,2,3,5,7,10,20 and 30-year) at the issuance date. Then, I interpolate them to build the yield of a Treasury bond that expires at the same time of the bond. For bonds with maturities larger than 30 years, I set the 30-year Treasury bond yield as the corresponding Treasury yield.

For each bond, I calculate a price in two ways. The first is directly, by just taking the recorded price at issuance. The second is indirectly, by projecting all the coupon payments (assuming yearly interest-only coupons and a single principal payment at maturity) and discounting them at the recorded yield at issuance. By definition, the price at issuance, coupons and yield at issuance of a bond are linked by

$$p = \sum_{t=1}^{T} \frac{c_t}{(1+y)^t} \tag{76}$$

where c_t is the coupon payment at time t (including principal and interest), T is the bond's maturity, y is the yield and p is the price. This implies that in theory the indirect calculation should give the same answer as directly recording the price, up to some inaccuracy in the exact timing of coupons, which the database does not detail. Indeed, for 90.3% of bonds, the two measures give answers within 1% of each other. However, there are some discrepancies in the database, often because the price-at-issuance is just recorded as equal to the face value. Because of this, the indirect calculation seems more reliable, and whenever I have information about the bond's yield at issuance I record the price using the indirect calculation; for bonds where the yield-at-issuance information is missing but I do have the issuance price I use the issuance price directly.

Since bonds differ across many dimensions, prices are not directly comparable across bonds. For instance, lower-coupon bonds will have a lower price than higher-coupon bonds of the same maturity and default probability. In order to have a measure of p that is comparable across bonds, I first compute the promised present value for each bond by discounting the projected coupons at the maturity-matched Treasury rate. I then normalize the price of each bond by dividing it by the promised present value. From this I obtain a measure of price per unit of promised present value.

The NYU Salomon database contains a listing of all bonds issued in the same 1977-2010 sample period that subsequently defaulted, including those that were originally issued as junk bonds and those that were not. I add a binary default indicator to each bond in the main sample that is also found in the NYU Salomon database. I match a bond in the main sample to one in the NYU Salomon database whenever (a) the entry in main sample includes the CUSIP indentifier and it matches an entry in the NYU Salomon database, (b) the entry in the main sample lacks a CUSIP identifier but the bond (i) is issued the same year, (ii) is issued be the same issuer and (iii) has either the same initial volume or the same coupon rate as an entry in the NYU Salomon database.

Investment bank profitability is measured as $\frac{\text{Net Income}}{\text{Net Worth}}$ for all firms in Compustat classified as investment banks. GDP growth is real GDP growth from NIPA. Stock market excess returns are the return on the S&P500 index minus the return on 3-month T-Bills. The price-earnings ratio is the cyclically adjusted price-earnings ratio computed by Robert Shiller (http://www.econ.yale.edu/~shiller/data.htm).