The Social Value of Financial Expertise

Pablo Kurlat*

Stanford University

December 2016

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. For junk bond underwriting I measure $r = 0.18$ and for venture capital I measure $r = 0.73$. In both cases this is less than one, which 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, Alejandro Martins, Thomas Philippon, Monika Piazzesi, Martin Schneider, Johannes Stroebel, Victoria Vanasco and Iván Werning for helpful comments. 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 remote from the production of goods and services, into 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 allocations and creating gains from trade. 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.

Formula (1) applies under the polar assumptions that the value of bad assets is zero and that households know perfectly which assets are bad. Before turning to empirical applications I first show how to extend the formula when these assumptions are relaxed. If the value of bad assets is positive, one just needs to measure it and incorporate it into a generalized version of formula (1). Accounting for the possibility that households are not perfectly informed 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 implement formula (1) empirically for two applications: junk bond underwriting and venture capital. For for junk bonds I rely on a combination of existing studies on underwriting fees (Datta et al. 1997, Jewell and Livingston 1998, Gande et al. 1999) and new estimates based on the universe of junk bonds issued around 1990, exploiting the variation around the collapse of the investment bank Drexel Burnham Lambert. I find $r = 0.18$. For venture capital, I rely on existing studies (Gompers and Lerner 2000, Hall and Woodward 2007) that break down the returns to venture investment by investor and over time, and find $r = 0.73$. These figures imply that of the last dollar earned by junk bond underwriters and venture capitalists by investing in expertise, 18 and 73 cents respectively is value added and the rest is captured rents. By these measures, both industries overinvest in expertise. There is considerable uncertainty around these numbers. The conclusion that $r < 1$ seems fairly robust for junk bond underwriters but less so for venture capital.

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 bonds and venture capital. Section 5 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.

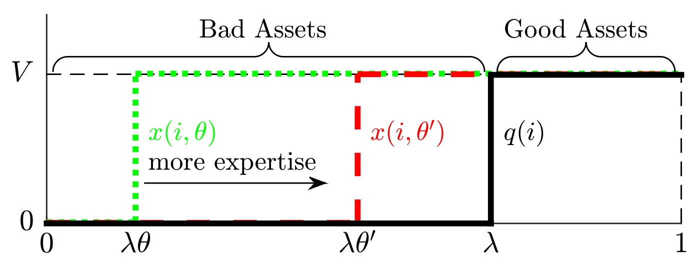

Figure 1: Asset qualities and signals

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.

[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.

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 the Appendix.

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 net payoff 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

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

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 net payoff 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:

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

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 the Appendix, 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:

$$S = (1-\lambda)\int_0^{s^*} (1-\beta(s))\,V ds \tag{10}$$

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

$$r_j \equiv \frac{S_j'(\theta_j)}{w(j)\,\tau'(\theta_j)} \tag{11}$$

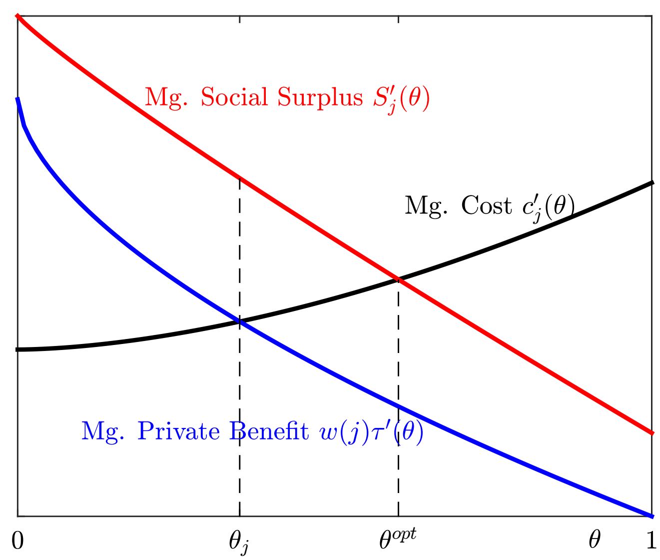

Figure 2: Example of marginal social surplus, private benefit and cost of additional investments in expertise. Bank j will choose expertise θ_j, equating marginal private benefit and marginal cost. The socially optimal level of expertise would be θ^{opt}.

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

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

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

$$\frac{\partial K_{3}^{*}}{\partial\theta}=-w_{j}\frac{\lambda}{\left[\left(1-\lambda\right)s^{*}+\lambda\left(1-\theta\right)\right]^{2}} \tag{19}$$

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:

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

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:

$$\alpha = \frac{(1-\lambda)\, s^* V}{\int_{\theta^*}^1 dW(\theta)} \tag{21}$$

The numerator represents the total dividends obtained from assets acquired by banks and the denominator is the total funds they spend.

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:

$$N = \frac{\int_{\theta^*}^1 dW(\theta)}{p^*} \tag{22}$$

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

The numerator in (22) is the total funds spent by banks who choose to trade and the denominator is the price they pay per asset. Therefore:

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

Notice that measuring f only requires tracking failures among assets that actually trade, not

[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.

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.

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

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

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 $\frac{ds^*}{d\theta}$. 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.

The quantities α and f can be measured relatively straightforwardly because they are simple averages. 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:

$$\eta = \frac{d\log(1-f)}{d\Delta} + \frac{d\log\left(\int_{\theta^*}^{1} dW(\theta)\right)}{d\Delta} - \frac{d\log(p^*)}{d\Delta} \tag{28}$$

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:

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

However, it is not unreasonable to assume that $w(\theta^*) = 0$. Choosing $\theta = \theta^*$ means that a bank would earn $\tau(\theta^*) = 0$ despite having invested a strictly positive amount of resources in acquiring expertise. Assuming $w(\theta^*) = 0$ 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.

$$\frac{d\log\left(\int_{\theta^*}^{1} dW\left(\theta\right)\right)}{d\Delta} = 1 \tag{29}$$

and therefore $d\Delta$ and be replaced with $d\log\left(\int_{\theta^*}^{1} dW\left(\theta\right)\right)$ in formulas (27) or (28).

3.3 Recovery Value

The baseline model makes the extreme assumption that the dividend from bad assets is exactly zero. While useful for theoretical clarity, in applications in may be desirable to relax this assumption. For instance, lenders typically recover a positive fraction of the value of loans that default.

Suppose that instead of zero, the dividend from bad assets was ϕV, where $\phi \in (0, 1)$. Under this assumption, there will be two markets (prices) with active trade in equilibrium. One of them would be similar to what happens in the baseline model: both good and bad assets will be on sale and only sufficiently expert banks will 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:

$$p^* = \frac{s^*\left(1-\lambda\right) + \lambda\left(1-\theta^*\right)\phi}{s^*\left(1-\lambda\right) + \lambda\left(1-\theta^*\right)} V \tag{30}$$

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 if it turns out to be bad.

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

$$S = \left[\left(1-\lambda\right)\int_0^{s^*}\left(1-\beta\left(s\right)\right)ds + \lambda\phi\int_0^{1}\left(1-\beta\left(s\right)\right)ds\right]V \tag{31}$$

The first term in (31) 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.[4] Using (31), the expression for the marginal social surplus is unchanged, still given by (13). The marginal private value of expertise is now:

$$w(j)\,\tau'(\theta_j) = \frac{w(j)}{p^*} V \frac{\lambda(1-\lambda)\,s^*(1-\phi)}{\left[(1-\lambda)\,s^* + \lambda(1-\theta_j)\right]^2} \tag{32}$$

Other things being equal, a higher recovery value lowers the marginal 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:

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

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 (32) and (33) 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.[5] The formula for α takes into account the dividends banks obtain from bad assets and therefore becomes:

$$\alpha = \frac{N\left[(1-f) + f\phi\right]}{\int_{\theta^*}^{1} dW(\theta)} \tag{34}$$

Formula (25) generalizes to

$$\frac{1-f+f\phi}{\alpha} = \beta(s^*) \tag{35}$$

and formula (26) for η becomes:

$$\eta = -\frac{1}{|D|}\frac{\lambda(1-\lambda)(1-\phi)}{\left[(1-\lambda)\,s^* + \lambda(1-\theta^*)\right]^2} V p^* \tag{36}$$

---

[4] If we had $\beta(s) > 1$ for some s this would no longer be true because a household with $\beta(s) \in \left(1, \frac{\phi V}{p^*}\right)$ 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 $p = \phi V$.

[5] 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.

Replacing (22), (24), (34) and (36) into (20) results in:

$$r = \frac{\eta}{\alpha}\left(\frac{\alpha - 1}{(1-\phi)} + f\right) \tag{37}$$

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.

3.4 Ex-Post Risk

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). Similarly, the measured rate of failed assets corresponds exactly to the fraction of traded assets that are bad.

Suppose instead that even good assets fail with probability π. Let $\hat{\phi}$ denote the measured recovery rate from assets that fail. The expected value of a good asset is $\pi\hat{\phi}V + (1-\pi)V$, so the relative value of bad assets is

$$\phi = \frac{\hat{\phi}}{\pi\hat{\phi} + (1-\pi)} \tag{38}$$

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:

$$f = 1 - \frac{1-\hat{f}}{1-\pi} \tag{39}$$

Replacing (38) and (39) into (37) to express r in terms of measurable quantities yields:

$$r = \frac{\eta}{\alpha}\left((\alpha - 1)\left(1 + \frac{\hat{\phi}}{(1-\pi)\left(1-\hat{\phi}\right)}\right) + 1 - \frac{1-\hat{f}}{1-\pi}\right) \tag{40}$$

which reduces to (37) 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 bias goes without assigning an precise value to π. Taking the derivative of (40):

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

This implies that if:

$$\Upsilon \equiv (\alpha - 1) \frac{\hat{\phi}}{1-\hat{\phi}} - \left(1 - \hat{f}\right) > 0 \tag{41}$$

then the underestimate of ϕ is more severe than the overestimate of f and the value that results from assuming $\pi = 0$ and applying (37) is a lower bound for r, and vice-versa.

4 Applications

4.1 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. After each underwriter has bought bonds, it can credibly disclose its information about them and re-sell them to investors who make zero profits. Inframarginal underwriters earn the profits indicated by equation (7) because they buy bonds at a fair price conditional on the marginal bank's information but re-sell them at a fair price conditional on their own information. I map these profits to the underwriting spread.

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 $\alpha = 1.046$.

In order to obtain measures of η and f, I focus on the period around 1990 in order to exploit the bankruptcy of the investment bank Drexel Burnham Lambert as a source of variation. I construct a sample that includes (subject to data availability) all the corporate bonds denominated in US dollars, issued between 1987 and 1990 and rated below investment grade by either S&P, Moody's or Fitch, a total of 585 individual bonds. The source is the Bloomberg database. For each bond I observe the total dollar amount issued, its coupon rate, its maturity, the yield spread against treasuries of comparable maturity and a binary indicator of whether it subsequently defaulted.[6]

For each bond j, I measure V_j by discounting its coupon and principal payments at the treasury rate of the corresponding maturity at the time of issuance. I then construct a measure of p_j by discounting the same coupon and principal payments at the bond's actual yield-at-issuance constructed by adding the bond's spread to the treasury rate. For each period t of the sample, I compute p_t as the dollar-weighted average of $\frac{p_j}{V_j}$, f_t as the dollar-weighted fraction of bonds issued in period t that subsequently default and $\int_{\theta^*}^{1} dW_t\left(\theta\right)$ by adding the dollar amount of all the bonds issued.

I measure f simply as the dollar-weighted fraction of bonds that defaulted in the whole sample, and obtain $f = 0.09$. This is somewhat lower than the numbers reported in previous studies (Altman 1989, 1992, Asquith et al. 1989, McDonald and Van de Gucht 1999, Zhou 2001), possibly as a result of unreported events of default.

I follow an event-study type of approach to measure η. The investment bank Drexel Burnham Lambert filed for bankruptcy in February 1990 following an SEC investigation for various forms of wrongdoing. Drexel was a major participant in the junk bond market, with a market share above 40%, and its demise had a major impact on the market (Brewer and Jackson 2000). I exploit the variation in volumes of bonds issued, bond prices and default

[6]I don't observe all of these measures for all the bonds. In particular, data on spreads is missing for many of them, so I exclude them from measures of p, though not from measures of total volume.

rates around 1990 in order to obtain a measure of η. To do this, I separately regress $\log(p_t)$ and $\log(1-f_t)$ on $\log\left(\int_{\theta^*}^{1} dW_t(\theta)\right)$ using yearly aggregates for 1987-1990. I find coefficients of 0.004 and 0.14 respectively. Applying formula (28) and assuming $w(\theta^*) = 0$ so that (29) holds, this results in $\eta = 1.13$.

The identification assumption is that around the window of Drexel's bankruptcy, the bankruptcy itself was the main source of movement in the junk bond market. There are good reasons to question this assumption. The economy was entering a recession at the time, and junk bond defaults were rising. It is possible that changes in the creditworthiness of issuers drove the variation in volume, prices and subsequent defaults. This would mean that both the elasticity of $1-f$ and the elasticity of p to capital flows are overestimated, so the net bias in the measurement of η (and therefore r) could go in either direction.

One alternative is to focus on a narrower window around Drexel's bankruptcy filing, where it is somewhat more likely that the bankruptcy itself is driving the variation. If I restrict the exercise to just two time periods: November 1989-January 1990 and March 1990-May 1990, I obtain $\eta = 1.28$. However, given the time it takes to arrange a bond issue, a too-narrow window is not ideal either.

The final parameter to measure is the recovery rate 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. Based on these studies, I set $\phi = 0.4$.

Replacing the measured values of $\alpha = 1.046$, $\eta = 1.13$, $f = 0.09$ and $\phi = 0.4$ into formulas (37) and (41) gives $r = 0.18$. This means that out of the last dollar that junk bond underwriters earn by being good at certifying the quality of bond issuers, 18 cents are value added and the remainder is captured rents. Furthermore, applying formula (41) gives $\Upsilon = -0.88$ and since this is below zero, this implies that $r = 0.18$ 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. Compared to the social optimum, the junk bond underwriting industry dedicates too many resources to the acquisition of expertise.

The main reason for the low value of r is that applying formula (35) gives $\beta(s^*) = 0.90$. This is a result of the low measured values of α and f. The social gains from the marginal trade are not very large, only 10% of the value of good assets. This could be because marginal junk bond issuers have alternative sources of funding (for instance, bank loans) or because they are close to indifferent between obtaining financing or not.

The measured values of α, f, ϕ and especially η are at best point estimates, in some cases derived from only a handful of observations, so the exact numbers should be treated with caution. Figure (3) shows how r changes with each of these variables. The range of values is not a confidence interval in a statistical sense, but is indicative of how sensitive the measured value of r is to each of its components. 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.

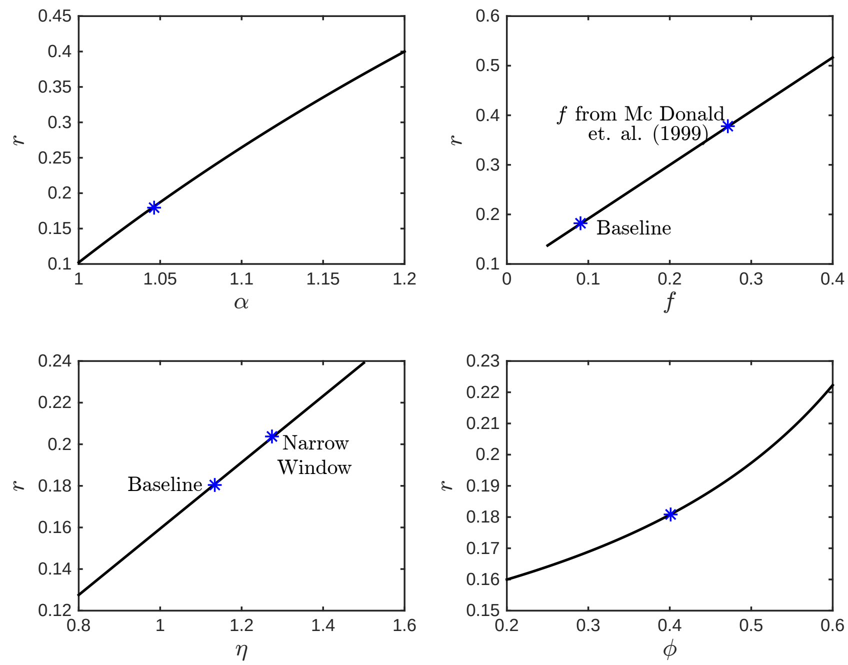

Figure 3: Sensitivity of r to each variable.

4.2 Venture Capital

The second application is more speculative, since within the literature on venture capital, there is some debate about whether asymmetric information is a major issue at all. Gompers (1995), Amit et al. (1998) and Ueda (2004) find evidence consistent with informational asymmetries, but the debate is not settled. I will assume that there is indeed asymmetric information, but the measure of r is conditional on the validity of this assumption.

Hall and Woodward (2007) use a large database of venture investments to measure how the value of venture-backed firms is, on average, split between the firm's founders and the general and limited partners of venture funds. I map these participants to the model as follows. The firm's founders are like the households in the model. They own an asset (the firm) and there are possible gains from trade in transferring part of the ownership of the firm to the venture fund. The general partners of venture funds are like the banks in the model. They have expertise in determining which firms are valuable. The limited partners are absent from the model. Hall and Woodward find that limited partners, who provide capital to venture funds but are not directly involved in decision-making, get almost no risk-adjusted excess returns from venture investments. I assume that general partners commit to deliver zero excess returns to limited partners and keep all excess returns for themselves in the form of fees. If this is true, the incentives to acquire expertise are proportional to the capital that the general partners administer. Hence $w(j)$ in the model corresponds to the total capital administered by a venture fund, including the capital supplied by limited partners.

Hall and Woodward find the the average net present value of all fees earned by general partners is 26% of the funds invested.[7] This suggests a value of $\alpha = 1.26$. This is probably an upper bound on α (and therefore an upper bound on r) since the rewards to venture capitalists compensate them for other services they provide firms besides screening them.

Gompers and Lerner (2000) use a different but overlapping database on venture investments to estimate the elasticity of valuations for venture-backed firms with respect to inflows of capital into venture funds. To do this they regress valuations of venture-backed firms at the time of a venture funding round, which I map to $\log p_t$, on the toal volume of funds committed to venture funds in a given time period, which I map to $\log\left(\int_{\theta^*}^1 dW_t(\theta)\right)$. They estimate a coefficient between 0.12 and 0.22 depending on the specification used. They don't report regressions with $\log(1-f_t)$ on the left hand side but it's possible to reconstruct them on the basis of the time series of f_t that they do report. Based on this data, the regression coefficient of $\log(1-f_t)$ on $\log\left(\int_{\theta^*}^1 dW_t(\theta)\right)$ is between 0.11 and 0.21 depending on the exact definition of a successful venture that is used. Using these values in formula (28) and assuming $w(\theta^*) = 0$ so that (29) holds gives a range of $\eta \in [0.89, 1.14]$. I'll take the midpoint of this range, $\eta = 1.01$, as the baseline figure.

Gompers and Lerner's estimates are based on exploiting time-series variation in inflows to venture funds. The identification assumption is that these inflows and outflows are exogenous, which is questionable. Possibly, funds flow into venture funds attracted by better

[7] These fees have two main components: a management fee that is usually a percentage of all funds committed and a "carry", set as a percentage of a fund's profits. The 26% figure is the average present value of the sum.

prospects for firms, which leads to higher prices and lower failure rates. Gompers and Lerner control for the most plausible channels of reverse-causality by including measures of stock market valuation as controls and by using inflows into leveraged buyout funds as instruments. Furthermore, they argue that regulatory changes like the clarification of the "prudent man" rule that allowed pension funds to invest in venture capital and changes in the capital gains tax rate account for much of the variation. Still, it's possible that the measured elasticities have omitted variable bias. This would bias both the elasticity of $1-f$ and the elasticity of p upwards, with an uncertain net effect on η.

Both Hall and Woodward and Gompers and Lerner propose measures of f, the fraction of failures among venture-backed firms.[8] Gompers and Lerner propose using the failure to either conduct an IPO or be acquired at twice the original valuation as a definition of failure (that definition is implicitly used in the measured elasticity above). Under this definition, in their data, $f = 0.66$. Hall and Woodward report similar figures. In their sample, the fraction of venture-backed firms that have not been acquired nor undergone an IPO is $f = 0.65$. Since these figures are based on observing truncated histories and some firms might conduct IPOs or be acquired later, this should probably be regarded as an upper bound on f.

According to Hall and Woodward, it is rare for venture-backed firms that are not acquired or undergo an IPO to return much value to investors. This would suggest that a recovery value of $\phi = 0$ might be a defensible assumption. However, some venture-backed firms continue as privately held firms and produce positive (though rarely large) dividends. I take $\phi = 0$ as the baseline case and examine how the numbers change with higher recovery values.

Replacing $\alpha = 1.26$, $\eta = 1.01$, $\phi = 0.65$ and $\phi = 0$ into formula (37) gives $r = 0.73$. This means that for the last dollar that general partners of venture funds earn by being good at selecting which firms to invest in, 73 cents are value added and the remainder is captured rents. Applying formula (41) gives$\Upsilon = -0.35$, so r is an upper bound relative to a model where good assets can also fail, which is undoubtedly an important issue in this application. If these figures are correct, the venture capital industry also dedicates too many resources to the acquisition of expertise relative to the social optimum. However, the measured wedge between social and private incentives is smaller than for junk bond underwriters.

The finding that r is below 1 is more fragile than for junk bonds. Figure (4) shows how

[8] Asset payoffs in the model are binary, either 0 or V. Payoffs from venture-backed firms are far from binary. Many fail and pay close to zero while among the successful ones there is a long right tail of extremely successful ones. This can be reconciled with the binary-payoff model by assuming that the value of successful firms is a random variable $\tilde{V}$ with expected value V. If we assume that entrepreneurs are not privately informed about the realization of $\tilde{V}$, then the fact that it's random makes no difference.

sensitive the measured value of r is to each of these variables. For instance, the combination of $\alpha = 1.3$, $\eta = 1.25$, $f = 0.7$ and $\phi = 0.12$, which is not far from the point estimates, leads to $r = 1$. Given the uncertainty in the measurements, it is hard to say conclusively that the industry overinvests in expertise.

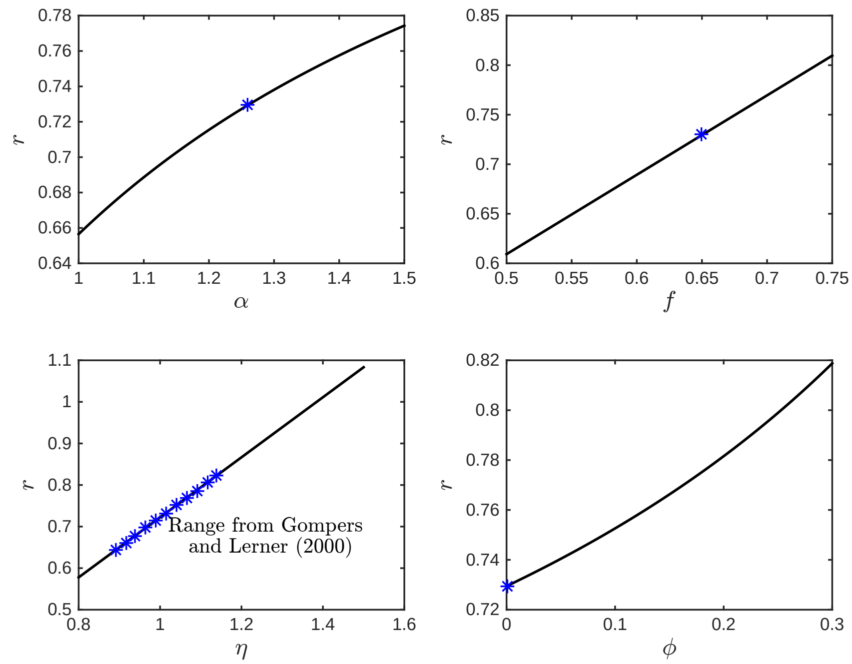

Figure 4: Sensitivity of r to each variable.

5 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 particular applications.

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}.

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.

The typical venture transaction differs from the simple outright sales that take place in the model: the venture capitalist's funds are invested in the firm rather that paid in cash to the founders. This is an important distinction but it need not change the basic force at play: venture capitalists demand a higher stake in the companies they finance than they otherwise would in order to compensate for investing in the firms that end up failing, and this discourages marginal entrepreneurs.

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 expands thanks to venture capital financing could generate positive externalities through technological spillovers or negative ones through business-stealing. 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.: 1989, Measuring corporate bond mortality and performance, *The Journal of Finance* **44**(4), 909–922.

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.

Amit, R., Brander, J. and Zott, C.: 1998, Why do venture capital firms exist? theory and canadian evidence, *Journal of business Venturing* **13**(6), 441–466.

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.

Asquith, P., Mullins, D. W. and Wolff, E. D.: 1989, Original issue high yield bonds: Aging analyses of defaults, exchanges, and calls, *The Journal of Finance* **44**(4), 923–952.

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.

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

Gompers, P. A.: 1995, Optimal investment, monitoring, and the staging of venture capital, *The Journal of Finance* **50**(5), 1461–1489.

Gompers, P. and Lerner, J.: 2000, Money chasing deals? the impact of fund inflows on private equity valuation, *Journal of financial economics* **55**(2), 281–325.

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.

Hall, R. E. and Woodward, S. E.: 2007, The incentives to start new companies: Evidence from venture capital. NBER working paper No. 13056.

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.

McDonald, C. G. and Van de Gucht, L. M.: 1999, High-yield bond default and call risks, *Review of Economics and Statistics* **81**(3), 409–419.

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.

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* .

Ueda, M.: 2004, Banks versus venture capital: Project evaluation, screening, and expropriation, *The Journal of Finance* **59**(2), 601–621.

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

Zhou, C.: 2001, Credit rating and corporate defaults, *The Journal of Fixed Income* **11**(3), 30–40.

Appendix

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.[9] 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.

A household of type s who owns asset i solves:

$$\max_{p^R} \quad \int_{p^R}^{V} p d\mu(p,i) + \left[1 - \mu\left(p^R, i\right)\right] \beta(s) q(i) \tag{42}$$

$$s.t. \quad \mu\left(p^R, i\right) \leq 1 \tag{43}$$

The first term in (42) 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 (43) limits the household to not sell more than one unit in total.

This problem as a simple solution. Define

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

$p^L(i)$ 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 $p^L(i) = 0$. It's immediate that the solution to program (42) is:

$$p^R(i,s) = \max\left\{p^L(i), \beta(s) V\right\} \tag{44}$$

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

[9] 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.

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 $\chi(i) = 1$ means that the bank is willing to accept asset i and $\chi(i) = 0$ 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 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. $\chi(i) = \chi(i')$ whenever $x(i,\theta) = x(i',\theta)$.

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 $A(\cdot;\chi,p)$ on the set of assets $[0,1]$ for each χ, p. For any subset $I \subseteq [0,1]$, $A(I;\chi,p)$ 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:

$$\max_{\delta,p,\chi} \quad \delta \left[\int_{[0,1]} q(i)\, dA(i;\chi,p) - pA([0,1];\chi,p) \right] \tag{45}$$

$$s.t. \quad \delta p A([0,1];\chi,p) \leq w \tag{46}$$

$$\chi(i) = \chi(i') \quad \text{whenever} \quad x(i,\theta) = x(i',\theta) \tag{47}$$

(45) adds all the dividends $q(i)$ of the assets the bank buys, subtracts what it pays per unit and multiplies by total demand δ; (46) is the budget constraint and (47) 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 (45) 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:

$$S(i;p) = \int_s \mathbb{I}\left(p^R(i,s) \leq p\right) \tag{48}$$

(48) 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:

$$D(X,p) = \int_{\theta \in \Theta(X,p)} \delta(\theta)\, dW(\theta) \tag{49}$$

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.[10] 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:

$$a(i;\chi,p) = \begin{cases} \frac{\chi(i)S(i;p)}{\int \chi(i)S(i;p)di} & \text{if } \int \chi(i)\, S(i;p)\, di > 0 \\ 0 & \text{otherwise} \end{cases} \tag{50}$$

Knowing A, the rationing faced by an asset i depends on the the ratio of the total demand

[10] 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.

that gets satisfied (added across all χ) to supply, so

$$\mu(p,i) = \int\limits_{\substack{\tilde{p}\geq p \\ \text{all } \chi}} \frac{a(i;\chi,\tilde{p})}{S(i;\tilde{p})} dD(\chi,\tilde{p}) \tag{51}$$

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; $\{\delta(\theta), p(\theta), \chi(\theta)\}$ 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 (50) and (51).

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 $\{p^R, \delta, p, \chi, \mu, A\}$ 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 $\{p^R, \delta, p, \chi, \mu, A\}$ is a conditional equilibrium given $W(\theta)$.

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:

$$p^R(i,s) = \begin{cases} \max\{p^*, \beta(s)V\} & \text{if } i \geq \lambda \\ 0 & \text{if } i < \lambda \end{cases} \tag{52}$$

2. The solution to the banks' problem is:

$$\{\delta(\theta), p(\theta), \chi(\theta)\} = \begin{cases} \left\{\frac{1}{p^*}, p^*, \mathbb{I}(i \geq \lambda\theta)\right\} & \text{if } \theta \geq \theta^* \\ \{0, 0, 0\} & \text{if } \theta < \theta^* \end{cases} \tag{53}$$

3. The allocation function is:

$$a(i; \chi, p) = \begin{cases} \frac{\beta^{-1}\left(\frac{p}{V}\right)\chi(i)}{\int_0^\lambda \chi(i)di + \int_\lambda^1 \chi(i)\beta^{-1}\left(\frac{p}{V}\right)di} & \text{if } i \geq \lambda \text{ and } p \geq p^* \\ \frac{\chi(i)}{\int_0^\lambda \chi(i)di + \int_\lambda^1 \chi(i)\beta^{-1}\left(\frac{p}{V}\right)di} & \text{if } i < \lambda \text{ and } p \geq p^* \\ 0 & \text{if } i \geq \lambda \text{ and } p < p^* \\ \frac{\chi(i)}{\int_0^\lambda \chi(i)di} & \text{if } i < \lambda \text{ and } p < p^* \end{cases} \tag{54}$$

4. The rationing function is:

$$\mu(p, i) = \begin{cases} 1 & \text{if } i \geq \lambda, p \leq p^* \\ \int_{\theta^*}^{\frac{i}{\lambda}} \frac{1}{\lambda(1-\theta)+s^*(1-\lambda)} \frac{1}{p^*} dW(\theta) & \text{if } i \in [\lambda\theta, \lambda), p \leq p^* \\ 0 & \text{if } i < \lambda\theta, p \leq p^* \\ 0 & \text{if } p > p^* \end{cases} \tag{55}$$

Proof.

(a) **Equations** (52)-(55) **constitute an equilibrium.**

i. Household optimization. (55) 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 (52) solves the household's problem.

ii. Bank optimization.

A. $\chi(\theta)$ is the optimal acceptance rule because, given (54), any other rule that satisfies (47) 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 $p > p^*$:

$$\begin{aligned}
\frac{1}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta^*)} &< \frac{1}{p^*}\frac{s^*}{s^*(1-\lambda)+\lambda(1-\theta^*)} \\
\frac{p^*}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{s^*} &< \frac{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta^*)}{s^*(1-\lambda)+\lambda(1-\theta^*)} \\
\frac{p^*}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{s^*} &< \frac{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta)}{s^*(1-\lambda)+\lambda(1-\theta)} \quad \text{for all } \theta \geq \theta^* \\
\frac{1}{p}\frac{\beta^{-1}\left(\frac{p}{V}\right)}{\beta^{-1}\left(\frac{p}{V}\right)(1-\lambda)+\lambda(1-\theta)} &< \frac{1}{p^*}\frac{s^*}{s^*(1-\lambda)+\lambda(1-\theta)} \quad \text{for all } \theta \geq \theta^*
\end{aligned} \tag{56}$$

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 (56) 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 (46) binds; for $\theta < \theta^*$ there is no $\chi(\theta)$ that satisfies (47) and leads to a positive value for the objective (45). Therefore $\delta(\theta)$ is optimal .

iii. Consistency of A and μ. Replacing reservation prices (52) into (48) and using this to replace $S(i;p)$ into (50) leads to (54). Adding up demand using(53) and (49) and replacing in (51) implies (55).

(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 (44) and formula (48) for supply imply that for $p \in \left[p^L(\lambda), p^*\right]$:

$$S(i,p) = \begin{cases} \beta^{-1}\left(\frac{p}{V}\right) & \text{if } i \geq \lambda \\ 1 & \text{if } i < \lambda \end{cases} \tag{57}$$

(57) implies that all banks with $\theta > \theta^*$ can attain $\tau(\theta) > 0$ by choosing p^*. By (56), 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 (50):

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

Using (51), this implies that

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

and therefore

$$\begin{aligned} \mu\left(p^L(\lambda), \lambda\right) &\geq \int\limits_{\theta^*}^{1} \frac{1}{\beta^{-1}\left(\frac{p(\theta)}{V}\right) + \lambda(1-\theta)} \frac{1}{p(\theta)} dW(\theta) \\ &\geq \int\limits_{\theta^*}^{1} \frac{1}{s^* + \lambda(1-\theta)} \frac{1}{p^*} dW(\theta) \\ &= 1 \end{aligned} \tag{58}$$

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 (58) 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 (50) and (51), we have

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

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.

□

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 (52), 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.