Federal Reserve Bank of Chicago

Bank Capital Regulation With and Without State-Contingent Penalties

By: David A. Marshall and Edward S. Prescott

WP 2000-10

Bank Capital Regulation With and Without State-Contingent Penalties

David A. Marshall and Edward S. Prescott*

September 1, 2000

Abstract

A moral hazard model with exogenous bank franchise value is used to analyze bank capital regulation. Banks choose their capital structure as well as the riskiness and mean of their portfolio. The portfolio mean is determined by the level of costly screening. Screening and portfolio risk are private information, so there are two dimensions to the moral hazard problem. Deposit insurance gives banks an incentive to hold less capital, and to choose a higher-risk, lower-mean portfolio. To mitigate these incentives, capital requirements with and without ex post fines are studied. We find an endogenous reverse mean-variance trade-off in banks' portfolios. Prudent banks choose high-screening, low-risk portfolios and are virtually self regulating. Imprudent banks choose low-screening, high-risk portfolios. Without state-contingent penalties, optimal capital regulations are often V-shaped in bank franchise value. Adding state-contingent regulation can, significantly lower capital requirements. Optimal state-contingent regulations are characterized by fines on extreme right-hand-tail returns.

*Authors' affiliations: Federal Reserve Bank of Chicago, Federal Reserve Bank of Richmond, respectively. The views expressed in this paper are those of the authors and are not necessarily the views of the Federal Reserve Bank of Chicago, the Federal Reserve Bank of Richmond, or the Federal Reserve System.

1 Introduction

This paper studies optimal bank capital regulation and state-contingent regulation in the context of a simple banking model. In particular, we ask whether capital regulation can be made more effective by augmenting it with state-contingent penalties, assessed on the basis of the bank's performance. In our model, banks are socially useful because they possess a special technology that allows them to issue demand deposits that provide liquidity services and to make investments with higher mean return than other assets available to households. The key choices of the bank are its capital structure, portfolio risk, and screening effort (which affects its portfolio mean). The problem of moral hazard, induced by deposit insurance, distorts all three of these decisions relative to the social optimum. These distortions are partially mitigated by the bank's concern for its *franchise value*, a proxy for the expected future value of the bank that is lost when the bank defaults. We consider how capital regulation, and capital regulation plus state-contingent fines, can reduce these distortions, and how these mechanisms ought to be implemented.

Our interest in the potential role for state contingency in bank capital regulation is motivated by developments in bank regulatory policy over the last 15 years. Both the 1988 Basle Accord and the Federal Deposit Insurance Corporation Reform Act of 1991 (FDICIA) make heavy use of ex ante capital regulation. The Basle Accord requires banks to hold capital in proportion to the riskiness of their assets. FDICIA allows the Federal Deposit Insurance Corporation (FDIC) to limit the activities of banks with low levels of capital, and it even allows the FDIC to shut a bank down if its capital level becomes critically impaired.

Since 1988, there has been increasing dissatisfaction with the Accord. It assesses risk on an asset-by-asset basis, rather than on a portfolio basis. Assets within the same risk class are treated the same, even if they have wildly different risks.[1] Furthermore, the development of securitization and financial derivatives allows banks to repackaging their assets in ways that lower their capital requirements without substantively altering their risks.[2]

In response, regulators have considered alternative regulatory regimes

[1]For example, the 1988 Accord assigns a zero risk weight to all OECD debt. As a consequence, during the Asian financial crises, Korean government debt was assigned the same risk weighting as U.S. government debt.

[2]See Jones (2000) and Mingo (2000) for examples.

that incorporate various degrees of state contingency. An example is the December 1995 amendment to the Basle Accord, known as the "internal models approach", that determines the capital requirements for large banks' market risk.[3] Under the internal models approach, banks report their own internal Value at Risk (VaR) estimates to their regulator. The regulator then uses these reports to determine how much capital the bank should hold to cover its market risk. Regulators backtest the reported VaRs to see if they are reliable and to ensure truthful reporting by the bank. Capital requirements may be increased if the reported VaRs underpredict the ex post frequency of tail events. This backtesting feature introduces an element of state-contingency into the regulatory process.

A more dramatic proposal for state contingent regulation is the Federal Reserve Board's Pre-Commitment proposal, put forth for public comment in June 1995.[4] Under this proposal, banks would set their own regulatory capital subject to a penalty if, ex post, this capital proved inadequate (again, as measured by the bank portfolio's ex-post performance). Most recently, a major revision to the Basle Accord is under consideration. One important feature under consideration would be to use bank internal risk rating systems to determine capital requirements for credit risk. As with the internal models approach for market risk, this feature requires self-reporting by banks. Some form of state-contingent incentive structure presumably would be required to insure truthful reporting and model accuracy.

In recent years, interest in state-contingent regulation among bank regulators has abated. This may be unfortunate, since the theoretical literature has argued that state-contingent regulation is often welfare-enhancing, as compared to ex ante regulation. Furthermore, any regulatory regime that requires banks to self-report must provide incentives for these reports to be accurate as well as to provide incentives for banks not to take too much risk. Despite the importance of these issues, there have been relatively few papers

[3]Bank regulators make the (somewhat arbitrary) distinction between market risk and credit risk. Market risk is the risk of a decline in bank asset values due to price movements in financial markets, while credit risk is the risk that borrowers may default. (The 1988 Basle Accords apply only to credit risk.) While this distinction has always been dubious, it is increasingly hard to justify in a world of over-the-counter derivatives, where a counterparty's creditworthiness is often determined by the performance of its securities portfolio. The near-default of Long Term Capital Management in September 1998 provides a pointed example.

[4]See Federal Register, Vol. 60, No. 142, pp. 38142-38144 and Kupiec and O'Brien (1995a,b).

that actually examine the role of state-contingency in bank regulation, so there are still many unanswered questions. For example, to what extent is welfare improved by moving from traditional types of bank regulation (such as minimal capital requirements) to state contingency? What would the optimal state-contingent regulation look like? Are there political reasons why this optimal regulation might not implementable in practice?

1.1 Summary of Results

We examine the optimal choices of an unregulated bank, a bank regulator with full information, and capital regulations both with and without state-contingent penalties when bank choice of risk and screening are private information. First, we find that default risk is decreasing and discontinuous in franchise value. In particular, there exists a threshold franchise value above which the unregulated bank's choices imply very low default risk. The choices of these high franchise value banks are very close to the socially optimal choices. In this sense, high franchise banks are virtually self-regulating. For banks with franchise value below this threshold value, however, default risk in the absence of regulation is very high. The key task of regulation is to control the incentives of these lower franchise banks.

The discontinuous shift in default risk that occurs at the threshold franchise value results from the distortions associated with deposit insurance. It has been well known since Merton (1977) that deposit insurance effectively gives the bank a put option on its assets, with strike price equal to its liability to depositors. The main concern of banks with franchise value below the threshold is to maximize the value of this put option by choosing maximal risk and zero capital. For banks above the threshold, the main concern is preserving franchise value by choosing a low default risk.

This first set of results is consistent with a good deal of empirical work on bank risk taking. Keeley (1990) argues lower franchise value leads to more risk taking. He argues that barriers to entry gave banks a high franchise value until the 1980's and that is why, despite deposit insurance, there were few failures until the deregulation of the 1980's. Demsetz, Saidenberg, and Strahan (1996) find that franchise value is negatively correlated with risk-taking in bank data from the 1990s. The evidence cited in Berger, Herring, and Szego (1995) suggests that a higher capital ratio is associated with less bank risk, although that this relationship is often weak. Boyd and Gertler (1994) provide evidence that during the U.S. banking crises of the late 1980's large banks took more risk than smaller banks. They argue that large banks were too big to fail, and thus the implicit insurance of their large deposits gave them more value from the deposit insurance put option.

Our second result is that screening effort (which increases the mean return in our model) is increasing in franchise value. Combined with our first result, this implies a sort of reverse mean-variance trade-off for banks: Those banks who take on higher risk (low franchise banks) tend to be those who also do less screening, and so have lower-mean portfolios, while those banks who invest in lower-risk portfolios (high franchise banks) have portfolios with a higher mean return. Furthermore, we find that the low franchise value banks (those below the threshold value) expend substantially less screening effort than socially optimal.

We obtain our second result because our moral hazard problem contains both a risk and screening dimension. The deposit insurance put option gives banks – in particular, low franchise ones – an incentive to take risk. But an increase in risk lower the probability of solvency, which in turn lowers the bank's marginal benefit of screening relative to that of society's. In our environment the main welfare cost of deposit insurance is not risk per se, but instead the negative effect of risk on screening.

These results are consistent with the limited empirical evidence on the interaction between bank loan quality and risk. Berger (1995) shows that banks with high capital and low portfolio risk tend to have higher returns on equity. This positive relationship is strongest for the highest risk banks (which would be the banks with the lowest franchise value in our framework). As noted by Gan (1999), this result appears to contradict the usual finance-theoretic intuition that lower portfolio risk be associated with lower profitability. However, this result is consistent with the behavior of our model. In addition, our results are consistent with evidence of thrift behavior during the savings and loan crisis. Benston and Carhill (1994) and White (1991) argue that thrifts that failed were those that invested in sectors with which they had little experience, such as commercial real estate. White (1991) notes that Texas thrifts did far more poorly in real estate lending at that time than Texas banks. He argues that the banks had a good deal of experience in this sort of real estate lending, while thrifts had only recently been permitted to engage in this sort of lending following deregulation. These patterns are consistent with insufficient or ineffective screening by the thrifts that ultimately failed.

Our third result concerns the optimal implementation of pure capital requirements. Many discussions of capital requirements focus on capital as a cushion to protect the bank insurance fund. In our model, there are no deadweight costs of failure resolution, so there is no advantage from "cushioning" the insurance fund. Rather, in our model capital requirements improve welfare by altering the incentives of low franchise-value banks. In particular, increased regulatory capital lowers the strike price of the deposit insurance put option, making moral hazard a less attractive choice for low franchise banks. In addition, capital has the obvious effect of directly reducing the probability of bankruptcy, with the attendant loss of future bank value.

Somewhat surprisingly, we find that the optimal regulatory capital level is non-monotonic in franchise value. The optimal capital requirement rises with franchise value, but only up to a point. Thereafter, optimal regulatory capital falls steeply with franchise value until the threshold franchise level, at which point the incentives of banks are well aligned with those of the regulator. For franchise values above this point, minimal capital requirements are unnecessary.

Our fourth set of results concerns the optimal way to augment capital requirements with ex post fines. Adding ex post fines to capital regulations can be effective in deterring risk-taking behavior and represents an improvement over pure capital regulation. In particular, when the optimal fine is implemented there is less of a need to impose suboptimally high capital requirements. Two types of fines arise as part of the optimal regulatory regime. First, fines are imposed on banks when they experience *high returns,* lowering capital requirements relative to a capital regulation regime. This type of fine schedule is robust in our numerical results. Fining the highest returns acts to deter excessive risk taking. These fines are assessed with very low probability when the bank chooses the optimal level of portfolio risk, but they are assessed with a much higher probability if banks deviate from this optimal risk level. Second, we find for some parameterizations that it is also optimal to fine banks that experience low profits. The purpose of this second fine is to induce banks to choose the optimal level of screening.

There are a number of issues that we do not address. First, our social welfare function is that typically used in the public finance literature: The welfare criterion is the utility of the representative household. In using this approach, however, we ignore any welfare costs associated with cross-sectional wealth redistribution. For example, when a bank chooses to exploit the safety net by taking on excess risk, wealth (in terms of expected value) is transferred to the offending bank from taxpaying households (if failure resolution is financed by taxes) or from other banks (if failure resolution is financed by assessments on the banking industry). Does this zero-sum transfer represent a social cost? Certainly, transfers from hard-working taxpayers to the Charles Keatings of the world would be regarded as ethically questionable by most people. Preventing such transfers may be a major motivation in designing bank regulation. In this paper, however, we do not consider this source of social welfare loss.

Second, we have placed the liquidity value of insured deposits in household's utility function. This simplifies our model, but renders it unsuitable for assessing mechanisms that relax the amount of deposit insurance. To address these questions, a more fundamental theory of demand deposits is needed, possibly a variation of the Diamond and Dybvig (1983) framework. (See, for example, Gorton and Winston (1999).) Third, in our model a bank is just a firm with access to a particular investment technology and with the right to issue insured deposits. Of course, banks actually function as financial intermediaries, and this role may lead to regulatory implications that we do not address. Finally, we do not address the agency problem between bank managers and bank shareholders. (See Dewatrapoint and Tirole (1993) and Besanko and Kanatas (1996).)

1.2 Literature

Our paper builds on the vast literature on bank regulation. The classic works on the effect of deposit insurance are Merton (1977) and Kareken and Wallace (1978). Both of these papers considered full information, complete market environments and showed that the deposit insurance gives banks incentives to take risk. There have been two strands to the subsequent literature. First, there is a large literature that examines the response of a bank to an exogenous change in capital regulations. The substantive implications of these papers depend on assumptions about bank preference orderings. Rather than summarizing this literature, we refer the interested reader to Calem and Rob (1999) and the references therein.

The second strand of the bank regulation literature explicitly focuses on the problem of private information. Private information is typically modelled either as unobserved characteristics or unobserved action ("moral hazard"). An early paper in this literature is Chan, Greenbaum, and Thakor (1992), which separately considered both types of private information. In the environment with unobserved characteristics, they found that actuarially fair insurance was inefficient. In Giammarino, Lewis, and Sappington (1993), both bank type and bank effort are unobserved. However, the distribution of the bank's return, is public information. Consequently, there is no welfare gain from regulation that is contingent on the bank's performance. Campbell, Chan, and Marino (1992) examine capital regulations in an unobserved characteristics model where a bank can choose a high or low return project with different risk characteristics. As in our work, they assume that deposits generate liquidity services.[5] They consider the value of having a self-interested regulator monitor the bank. Boyd, Chang, and Smith (1998) use the costly state verification paradigm. Their model is general equilibrium and focuses on the best way to fund potential losses to the deposit insurer. In contrast with much of the literature, their paper explicitly models the role of banks as intermediaries.

Three other recent papers are closely related to our work. Nagarajan and Sealy (1998) considers state-contingent regulation. They find that ex post state-contingent fines can support full-information allocations in a moral hazard model and in an adverse selection model (though not both simultaneously). In their moral hazard model, banks control the characteristics of the return distribution. The optimal regulatory mechanism focuses on relative performance. That is, banks are fined when the bank does poorly but the market does well. Unlike our model, they do not have a well-defined social welfare function or a theory of capital structure determination. Besanko and Kanatas (1996) is perhaps closest to our paper in that they incorporate effort that affects returns but generates disutility. Unlike our work, they model bank insiders who raise capital from outside investors. As in their paper, a suboptimally low level of effort is supplied in the unregulated case. Our paper builds on their insights by incorporating a formal welfare criterion, introducing a liquidity value for deposits (which implies a wedge between the cost of equity and deposit funding), considering alternative bank franchise values, and allowing for a richer distribution of returns. In addition, we address the potential welfare gains from state-contingent regulation. One final relevant work is a recent paper by Matutes and Vives (2000), who consider a moral hazard problem in an economy where banks have market power. Market power provides an incentive for banks to avoid risk, similar to the role that franchise value plays in our model. They then study the trade-offs between the social costs to bankruptcy and the welfare losses from market power.

[5] Gorton and Winton (1999) also make this assumption.

They do not study ex post state-contingent regulatory mechanisms.

1.3 Structure of Paper

In section 2 of the paper, we describe our model of banking. In section 3, we characterize the optimal choices of the unregulated bank and the full-information regulator. Section 4 describes the regulatory regimes we consider. Section 5 describes the results from our numerical exercises for the unregulated bank, the full-information regulator, and pure capital structure regulation. In section 6, we discuss our numerical results for capital regulation with state-contingent incentives. Section 7 summarizes our conclusions and suggests directions for future research.

2 A Simple Static Model of Banking

2.1 Households

There are two periods and a single consumption good. There is a large number of identical risk-neutral households, who consume in the second period only. The households own all assets in the economy and consume all output. In the first period, the household receives an endowment of one unit of the consumption good. The head of the household, who we call the "banker", then decides how much of this endowment to use as bank capital K. The remainder of the endowment, $1 - K$, is then put in some other bank as a demand deposit. We assume that households cannot deposit funds with the banker in their own household. Demand deposits are government insured, and pay off one unit of the consumption good in the second period for each unit invested in the first period. In addition to its pecuniary payoff, a unit of bank deposits provides liquidity services with utility value $\rho > 0$. Parameter ρ can be interpreted as the bank's cost of capital relative to deposit financing.

The banker then takes in insured deposits totalling D from other households, and invests the total $K + D$ in a portfolio of assets. Since the endowment per bank is unity, a symmetric allocation has

$$K + D = 1. \qquad (1)$$

We impose equation (1) as a constraint of the model. That is, the representative bank's portfolio is limited to be of size one. One can think of this assumption as an extreme version of the limits imposed by managerial span of control.

The banker then constructs the bank's portfolio. For simplicity, we do not model the individual assets in the portfolio. Rather, we assume that the portfolio return, denoted r, has a log normal distribution with mean μ and variance σ^2. The banker chooses σ directly, subject to the constraint that $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ (where $\sigma_{\min}$ and $\sigma_{\max}$ are positive parameters). The mean of the bank's portfolio is determined by σ and by the banker's choice of screening effort $S \geq 0$ to use in evaluating alternative assets for the bank's portfolio. This screening effort has utility cost γS, with parameter $\gamma > 0$. The mean of the portfolio is a function $\mu(S, \sigma)$. It is assumed that $\mu(S, \sigma)$ is bounded, and that

$$\mu_S \geq 0, \quad \mu_\sigma \geq 0,$$

$$\mu_{SS} \leq 0, \quad \mu_{\sigma\sigma} \leq 0,$$

$$\mu_{S\sigma} \leq 0.$$

The assumption that $\mu_{S\sigma} \leq 0$ captures the idea that riskier projects are harder to screen. The assumption $\mu_\sigma \geq 0$ incorporates a non-negative relationship between risk and mean return. We allow $\mu_\sigma > 0$ for completeness. If all assets were market traded, $\mu_\sigma > 0$ would be incompatible with risk neutrality. However, bank portfolios include non-market traded assets such as real estate and commercial and industrial loans. There may be technological reasons why high mean assets of these types are associated with greater risk. Furthermore, allowing $\mu_\sigma > 0$ permits a greater variety of results. For example, $\mu_\sigma = 0$ implies that the bank's choice of σ is always either $\sigma_{\min}$ or $\sigma_{\max}$, while $\mu_\sigma > 0$ allows for interior choices of σ. In our numerical experiments we take $\mu_\sigma = 0$ as our baseline case. In our numerical simulations, we assume the following functional form for $\mu(S, \sigma)$:

$$\mu(S, \sigma) = \mu_{\max} - (\mu_{\max} - \mu_{\min}) \exp\left[-a_S S - a_\sigma \sigma\right], \tag{2}$$

where $\mu_{\max}$, $\mu_{\min}$, a_S, and a_σ are positive parameters with $\mu_{\max} \geq \mu_{\min}$.

Let $F(\cdot|S, \sigma)$ and $f(\cdot|S, \sigma)$ denote the log normal cdf and log normal density with mean $\mu(S, \sigma)$ and variance σ^2. The bank is subject to limited liability, so the expected bank payoff in the second period, v, is

$$v \equiv \int_D^\infty [r - D] f(r|S,\sigma)\, dr. \tag{3}$$

The household purchases consumption in the second period using its bank deposits plus the profits from the banker's activities, less a lump sum tax, with expected value T, used by the deposit insurer to pay off the depositors of failed banks. This mean tax level is given by

$$T = \int_0^D [D - r] f(r|S,\sigma)\, dr. \tag{4}$$

Therefore, the expected household consumption C is subject to the constraint

$$C \leq D + v - T \tag{5}$$

In addition, we associate with the bank a franchise value ψ (in units of consumption good). Franchise value stands for the future value of the bank as an ongoing institution. This franchise value is lost (both to the bank and to society) if the bank defaults on its obligations to depositors. That is, ψ cannot be sold to satisfy claims of the bank's liability holders. We think of ψ as representing the portion of bank value tied up in its pre-established relationships with borrowers.[6]

Household utility is a linear function of consumption, liquidity services provided by deposits, the disutility. of screening effort, and the expected franchise value. Since franchise value is lost if the bank defaults, an event with probability $F(D|S,\sigma)$, the problem faced by the banker is:

$$\max_{C,D,S,\sigma} C + \rho D - \gamma S + (1 - F(D|S,\sigma))\,\psi \tag{6}$$

subject to equations (3), (5), and

$$0 \leq D \leq 1, \sigma_{\min} \leq \sigma \leq \sigma_{\max}, S \geq 0. \tag{7}$$

Note that equation (4) is an equilibrium condition so the bank takes taxes T as exogenous when solving problem (6).

We can re-write the bank's problem in perhaps a more intuitive way. Using equation (3), we can write bank problem (6) as follows:

[6] Diamond and Rajan (1998) discuss at length the theoretical implications of non-tranferability of the value of bank-borrower relationships.

Unregulated Bank Problem

$$\max_{D,S,\sigma} \int_0^D (D-r)\, f(r|S,\sigma)\, dr + \rho D + \mu(S,\sigma) - \psi F(D|S,\sigma) - \gamma S + \psi - T \quad (8)$$

subject to inequalities (7). The first term in equation (8) is the payoff of the deposit insurance put option with strike price D. The second term is the value of liquidity services received from deposits. The third term is the mean return to the bank's portfolio. The fourth term is the expected loss in franchise value due to the possibility of default, and the fifth term is the lost utility due to screening activity.

To solve for an equilibrium, we solve problem (6) (or problem (8)) subject to (7) and treat (4) as an equilibrium condition.

2.2 Regulator

We make the assumption that the regulator seeks to maximize the utility of the representative household, as given in the maximand of equation (6), subject to equations (3), (4), (5), and (7).[7] The difference between the regulator's problem and the bank's problem is that the bank takes the level of taxes as exogenous, while the regulator takes into consideration the way banks' choices determine the lump-sum tax.

Suppose the regulator could observe and control the bank's choices of $\{S, D, \sigma\}$. Substituting equations (3), (4), and (5) into equation (6), one obtains:

Full Information Regulator's Problem

$$\max_{S,D,\sigma} \rho D + \mu(S,\sigma) - \psi F(D|S,\sigma) - \gamma S + \psi \quad (9)$$

subject to inequalities (7). Note that, except for constant terms unaffected by the choice variables (including T in the bank's problem), the only difference between problems (8) and (9) is that problem (8) includes the payoff to the

[7] Of course, regulators objectives need not be the same as society's, in which case agency problems may arise between regulators and taxpayers. We abstract from these important issues.

deposit insurance put option. Other than this option component, the bank's and regulator's incentives are identical. Therefore, this model focuses exclusively on welfare distortions from moral hazard induced by deposit insurance. Other potential distortions that regulators might be concerned about include deadweight costs of failure resolution, distortionary taxation that might be used to pay off depositors of failed banks, welfare losses due to wealth redistribution from taxpayers to bankers exploiting the safety net, and systemic crises that might be triggered by widespread bank failures. Our model does not address these potentially important issues.

2.3 Digression on actuarially fair deposit insurance premiums

A comparison of problems (9) and (8) suggests that the regulatory optimum could be achieved if the regulator were to impose an insurance premium that varies with D, S, and σ, according to

$$\int_0^D (D - r)\, f\,(r|S, \sigma)\, dr. \tag{10}$$

To do so, however, would require the regulator to observe the bank's choice of S and σ. In our discussion of regulatory mechanisms, we assume that S and σ are private information. Therefore, a "risk-based" deposit insurance premium of this type is not feasible. Note that the optimum cannot be achieved by imposing a premium equal to the object in (10) evaluated at the *regulator's* optimal choice of $\{S, D, \sigma\}$, nor would it be achieved by a premium equal to (10) evaluated at the bank's choices under the insurance premium regime. In both of these cases the bank regards the premium as fixed, so it does not offset the convex payoff structure introduced by the term $\int_0^D (D - r)\, f\,(r|S, \sigma)\, dr$ in problem (8).

3 Characterizing the Bank and Regulatory Optima

3.1 Unregulated Bank's Problem

Let S^*, D^* , and σ^* denote the solution to problem (8). The following proposition characterizes these choices of the unregulated bank.

Proposition 1

1. *$D^* > 0$*
2. *If $\psi = 0$, $D^* = 1$.*
3. *Suppose $0 < D^* < 1$. Holding S^* and σ^* fixed,*

$$\frac{\partial D^*}{\partial \psi} < 0 \text{ and } \frac{\partial D^*}{\partial \rho} > 0$$

4. *If μ_s is sufficiently small, $S^* = 0$.*
5. *Suppose $S^* > 0$. Holding D^* and σ^* fixed,*

$$\frac{\partial S^*}{\partial \psi} > 0, \frac{\partial S^*}{\partial \gamma} < 0$$

(Proof in Appendix A).

Part 1 of proposition 1 says that the unregulated bank never chooses 100% capital financing. Part 2 states that a zero franchise value bank always chooses zero capital. By continuity, we can then infer that this result also holds for banks with franchise value sufficiently low. Part 3 of proposition 1 has a simple interpretation: As the value of continued survival increases, banks choose to hold more capital (lower D^*); but as the value of liquidity services increases (higher ρ), banks hold less capital. The interpretation of part 5 is similar. Increasing S has two salutary effects on the maximand in problem (8): it increases $\mu(S, \sigma)$ and it decreases $\psi F(D|S, \sigma)$. As ψ increases, this second effect becomes increasingly important, increasing the equilibrium choice of S.

In parts 3 and 5 of proposition 1 we sign the partial derivatives of D^* and S^* with respect to parameters of the model. The total derivatives cannot be signed because the sign of the cross-partial derivative with respect to $\{D, S\}$ of the maximand in problem (8) is ambiguous. In particular, let $F^{\ln}(x|\mu, \sigma)$ and $f^{\ln}(x|\mu, \sigma)$denote the log normal cumulative distribution function and probability density function with mean μ and variance σ^2. Note that $F^{\ln}$ is related to F by

$$F(x|S, \sigma) = F^{\ln}(x|\mu(S, \sigma), \sigma),$$

where $\mu(S,\sigma)$ is the screening function. The cross-partial derivative with respect to $\{D,S\}$ of the maximand in problem (8) is

$$\mu_S\left[F_\mu^{\ln}(D|\mu,\sigma)-\psi f_\mu^{\ln}(D|\mu,\sigma)\right]$$

where we use

$$F_S(D|S,\sigma)=F_\mu^{\ln}(D|\mu,\sigma)\mu_S.$$

According to lemma 4 in appendix A, $F_\mu^{\ln}(D|\mu,\sigma)<0$. However, $f_\mu^{\ln}(D|\mu,\sigma)$ cannot be unambiguously signed. (A similar ambiguity would arise if we were to assume normality, in which case $f_\mu(D^*|S,\sigma)<0$.) Having said this, we almost always find in our numerical simulations that the total effect of increasing ψ has the same sign as the partial effects described in proposition 1.

Proposition 1 does not discuss the behavior of σ^*. This choice variable is somewhat more difficult to analyze because it affects both the variance and (via the screening function $\mu(S,\sigma)$) the mean of the bank's portfolio return. As a result, F_σ cannot be unambiguously signed. However, we can get some interesting analytic results for the case where both μ_σ and $\mu_{\sigma\sigma}$ are sufficiently small. For the parameterization of $\mu(S,\sigma)$ we use, one can make both $\max[\mu_\sigma]$ and $\max|\mu_{\sigma\sigma}|$small by choosing a sufficiently small value for parameter a_σ. In proposition 2, we show that if $\max[\mu_\sigma]$ and $\max|\mu_{\sigma\sigma}|$ are sufficiently small, σ^* is never interior. Rather, the bank's optimal choice of risk is either $\sigma_{\min}$ or $\sigma_{\max}$. Furthermore, as franchise value increases, there comes a point at which the bank's optimal choice of risk shifts discontinuously from $\sigma_{\min}$ to $\sigma_{\max}$.

Proposition 2 *For* $\max[\mu_\sigma]$ *and* $\max|\mu_{\sigma\sigma}|$ *sufficiently small,*

1. $\sigma^*=\sigma_{\min}$ *or* $\sigma_{\max}$.
2. *If* $\psi=0,\sigma^*=\sigma_{\max}$.
3. *Fix* $\{S^*,D^*\}$. *There exists* $\overline{\psi}$ *such that* $\sigma^*=\sigma_{\max},\forall\psi<\overline{\psi}$, *and* $\sigma^*=\sigma_{\min},\forall\psi>\overline{\psi}$.

(Proof in Appendix A).

Proposition 2.1 implies that when μ_σ is low, banks endogenously sort into risk-seeking banks (low franchise value banks who choose $\sigma^*=\sigma_{\max}$)

and prudent banks (high franchise value banks who choose $\sigma^* = \sigma_{\min}$). For higher values of μ_σ, we find in our numerical simulations that interior values of σ do emerge in equilibrium. However, the distinction between risk-seeking vs. prudent banks continues to be useful as a qualitative description of bank behavior. We find that, for sufficiently low franchise values, banks set $\sigma^* = \sigma_{\max}$ and $D^* = 1$ (i.e., zero capital) in an effort to maximize the value of the deposit insurance put option. For higher franchise values, the advantage of exploiting this put option are outweighed by the loss in franchise value, so banks reduce their probability of failure by reducing σ^* and/or D^*discontinuously.

For fixed D^* and S^* Proposition 2.3 gives a comparative static result for σ^* comparable to those given in proposition 1 for D^* and σ^*: σ^* is (weakly) decreasing in ψ. For sufficiently small μ_σ these two results suggest a sort of "reverse mean-variance trade-off": As ψ increases, σ^* decreases while $\mu(S^*, \sigma^*)$ increases.

3.2 Regulator's problem under full information

Let D^{reg}, S^{reg}, and σ^{reg} denote the solutions to problem (9). These optimal choices of the regulator under full information are characterized by the following proposition.

Proposition 3

1. $D^{reg} > 0$.
2. *If* D^{reg} *is interior,* $\frac{\partial D^{reg}}{\partial \psi} < 0$ *and* $\frac{\partial D^{reg}}{\partial \rho} > 0$
3. *If* σ^{reg} *is interior and* μ_σ *is sufficiently small,* $\frac{\partial \sigma^{reg}}{\partial \psi} < 0$
4. *If* $S^{reg} > 0$*, then* $\frac{\partial S^{reg}}{\partial \psi} > 0$ *and* $\frac{\partial S^{reg}}{\partial \gamma} < 0$.

(Proof in Appendix A).

The interpretation of proposition 3 is analogous to that of proposition 1 for the unregulated bank. For the regulator, there is no result analogous to proposition 2.1, since there is no tension between the risk of lost franchise value and the gain from enhancing the value of the deposit insurance put option. In fact, the only reason that σ^{reg} might exceed $\sigma_{\min}$ is the mean-enhancing effect of higher variance through a positive value of μ_σ.

4 Regulatory Mechanisms

In the previous section we saw that the bank's objective function (8) differs from the objective of the regulator (9). As a result, the choices of the unregulated bank will generally be suboptimal from the perspective of the regulator. If the bank's choices of $\{S, D, \sigma\}$ were observable to the regulator, the solution to problem (9) could be implemented. However, a crucial difficulty faced by makers of regulatory policy is that the distribution of bank portfolio returns is imperfectly observable. As discussed in the introduction, the Basle Committee is currently struggling with the inaccuracy of its 1988 "buckets" approach to measuring bank portfolio risk. To capture this fundamental unobservability of bank portfolio risk, we assume that the bank's choice of $\{S, \sigma\}$ is private information, unavailable to the regulators. The bank's choice of D, however, corresponds to the bank's capital structure, which clearly is public information.

We further make the assumption that the realized return r to the bank's portfolio is observable. (In section 6.2, below, we will partially relax this assumption.) In the case of the bank's trading portfolio, this assumption approximates reality. For example, both the pre-commitment proposal and the backtesting feature of the internal models approach rely on the observability of realized returns to the bank's portfolio of marketed securities. The return to the bank's loan portfolio is only imperfectly observed by regulators. In assuming observability of r, we are assuming that bank examiners can form a fairly good estimate of the current value of a bank's loan book, even if they cannot accurately assess the probability distribution of its future value.

In this section we consider two regulatory mechanisms: Bank capital regulation, and optimal state-contingent regulation.

4.1 Bank Capital Regulation

In the context of our model, bank capital regulation corresponds to having the regulator choose a value for deposits $D^{cap} \in [0, 1]$, and letting the bank then solve problem (8) subject to the constraint that $D \leq D^{cap}$. We denote the optimal choices of the bank under this regulatory regime by $\{S^{cap}, \sigma^{cap}\}$.

A formal statement of the problem is as follows:

Bank Capital Regulation Problem:

$$\max_{S,D,\sigma,D^{cap}\in[0,1]} \rho D + \mu\left(S,\sigma\right) - \psi F\left(D|S,\sigma\right) - \gamma S + \psi \tag{11}$$

subject to equation (4), and the incentive constraints that

$$\{D,S,\sigma\} = \arg\max_{D\leq D^{cap},S,\sigma} \int_0^D \left(D-r\right) f\left(r|S,\sigma\right) dr + \rho D + \mu\left(S,\sigma\right) - \psi F\left(D|S,\sigma\right) - \gamma S + \psi - T$$

where $\{D,S,\sigma\}$ are subject to inequalities (7).

4.2 Optimal State-Contingent Regulation

We assume that the regulator can observe two variables correlated with the bank's choices: D and r. Bank capital regulation only exploits the observability of D. The optimal regulatory mechanism exploits observability of both of these variables. We model this mechanism as a fine $g(D,r)$, which is a function of the bank's choice of D and of the realized return r. The fine is assessed against the bank *ex post.* Regulators collect the fine, pay a resource cost of τ per unit of collected fine, and rebate the remainder to households as a reduction in their lump-sum tax. We impose the resource cost because the spirit of our analysis is to use fines to affect bank behavior, not to use regulators as a device for transferring wealth to households. Besides, it is reasonable to assume that there is some administrative cost to collect fines.

We impose two restrictions on the fine schedules. First, we require fines to be non-negative. We make this assumption because our theory of capital structure is based on insolvency and transfers from the FDIC could always be made to keep the bank solvent. Second, we require fines to satisfy a limited liability constraint.[8] More formally, fines are bounded by

$$0 \leq g(D,r) \leq \max\{0, r-D\}. \tag{12}$$

The problem for the regulator is now

Optimal State-Contingent Regulation Problem:

[8] One also might want to impose an upper bound on fines. For some distributions, with risk neutrality and a continuum of returns, no solution may exist because the algorithm would try to impose higher and higher fines on a smaller and smaller set of the returns. (See Mirrlees, 1999). This is not a problem for our numerical exercises, since we discretize the space of possible returns.

$$\max_{D,S,\sigma,g(\cdot,\cdot)} \rho D+\mu\left(S,\sigma\right)-\gamma S-\psi F\left(D|S,\sigma\right)+\psi-\int_0^\infty \tau g(D,r) f\left(r|S,\sigma\right) dr \quad (13)$$

subject to equation (12), to the incentive constraints

$$\begin{aligned}\{D,S,\sigma\} &= \arg\max_{D,S,\sigma} \int_0^D \left(D-r\right) f\left(r|S,\sigma\right) dr + \rho D + \mu\left(S,\sigma\right) \\ &\quad -\gamma S - \psi F\left(D|S,\sigma\right) + \psi - \int_D^\infty g(D,r) f\left(r|S,\sigma\right) dr - T,\end{aligned}$$

where $\{D,S,\sigma\}$ are subject to inequalities (7), and where taxes are now $T = \int_0^D \left[D-r\right] f\left(r|S,\sigma\right) dr - \int_0^\infty (1-\tau) g(D,r) f\left(r|S,\sigma\right) dr$. The last term in (13) represents the resources used to collect the fines. Finally, note that specifying the fine as a function of $\{D,r\}$ is equivalent to having the fine defined only over r with the regulator in addition choosing the bank's capital level. For this reason, we sometimes refer to this regulatory regime as capital regulation augmented with state contingent penalties.

5 Results from Numerical Solution of the Model: Unregulated Bank, Full-Information Regulator, and Pure Capital Regulation

5.1 Baseline Parameterization

In this subsection, we describe results from numerically solving problems (8), (9), and (11) for a variety of parameterizations. We defer our numerical analysis of problem (13) until section 6.

We start with the following baseline parameterization:

- $\rho = 0.05$
- $\gamma = 1$
- $\mu_{\min} = 1.05$
- $\mu_{\max} = 1.75$

- $\sigma_{\min} = 0.3$
- $\sigma_{\max} = 1.5$
- $a_s = 5$
- $a_\sigma = 0$

We think of the time between period 1 and period 2 as one year, so the 5% cost-of-funding advantage of deposits over capital implied by our choice of ρ seems plausible. With $\mu(S, \sigma)$ given by equation (2), parameter γ is essentially a normalization.[9] Parameter $\mu_{\min}$ is the portfolio mean that would obtain if there were zero risk and zero screening, so it is natural to set it to a value approximating the risk-free rate. Our choices for $\mu_{\max}$, $\sigma_{\min}$, and $\sigma_{\max}$ are somewhat more arbitrary. They imply that no portfolio can have a mean excess return above 70% or a standard deviation above 150%, and all portfolios have a standard deviation of at least 30%. These values are perhaps somewhat high (especially the specification for $\sigma_{\min}$) but they are convenient for illustrating the properties of the model. We do consider a variety of alternative specifications for these parameters. We have no clear guide for setting a_s so we start with a value of 5, and then consider alternative specifications. Finally, as discussed earlier, setting $a_\sigma = 0$ is a natural starting point in a model with risk-neutral agents.

Figure 1 plots $\{D^*, D^{reg}, D^{cap}\}$ (panel A), $\{S^*, S^{reg}, S^{cap}\}$ (panel B), and $\{\sigma^*, \sigma^{reg}, \sigma^{cap}\}$ (panel C) as functions of $\psi \in [0, 1]$ for the baseline model. In each of these panels, the solid line gives the optimal choices of the unregulated bank (i.e., the solution to problem (8)), the dashed line gives the optimal choices of the regulator with full information (the solution to problem (9)), and the dot-dashed line gives the equilibrium values under optimal capital regulation (that is, the values associated with the solution to problem (11)). In addition, we report the equilibrium values for the probability of default (panel D), the mean of the bank's portfolio (panel E), and the value of the social welfare criterion, given by the maximand of problem (9) (panel F).

Consider first a comparison of the unregulated bank's choices with the full-information regulatory optimum. There are five points in particular to note from Figure 1. First, the way choice variables $\{S^*, D^*, \sigma^*\}$ and

[9] Consider equation (2). If both γ and a_s were doubled, the equilibrium value of S would be halved, but all allocations and utility values would remain unchanged.

$\{S^{reg}, D^{reg}, \sigma^{reg}\}$ change as ψ increases correspond to the partial derivatives discussed in propositions 1, 2, and 3. That is, the signs of the total derivatives of these choice variables with respect to ψ are the same as those of the partial derivatives. In particular, the bank's and regulator's choice of screening is everywhere increasing in ψ, even though, for the higher ψ's, the rate of increase is very small (imperceptible from the graph).

Second, there is a threshold franchise value $\overline{\psi}$ (approximately 0.425 in this parameterization) at which bank risk-seeking shifts discontinuously to prudent behavior. At this franchise value, not only does σ^* shift from $\sigma_{\max}$ to $\sigma_{\min}$, but we also see a jump in S^*. These effects induce a discontinuous drop in the probability of default from 0.42 to near zero.

Third, for $\psi < \overline{\psi}$, there is a potential loss in social welfare from the absence of bank regulation. In particular, the value of the regulator's objective function evaluated at $\{S^*, D^*, \sigma^*\}$ (solid line in panel F) is substantially below that evaluated at $\{S^{reg}, D^{reg}, \sigma^{reg}\}$. Furthermore, this gap increases as ψ increases from zero to $\overline{\psi}$. As ψ increases, the bank is more valuable to society, yet (for $\psi < \overline{\psi}$) the actions of the unregulated bank do not protect this value sufficiently against default.

Fourth, for $\psi > \overline{\psi}$ the choices of the unregulated bank and the socially optimal choices are virtually identical. Numerically, D^* is still slightly above D^{reg} and S^* is slightly below S^{reg}, but these discrepancies are negligible. In this sense, high franchise value banks are virtually self-regulating. The only role for regulation is to affect incentives for the low franchise value banks.

Fifth, we find a sort of reverse mean-variance trade-off in equilibrium, since σ^* is (weakly) decreasing in ψ, while $\mu(S^*, \sigma^*)$ is strictly increasing in ψ. This, if we observed an economy with a distribution of franchise values, we would find that higher-risk banks have portfolios with lower mean returns.

Let us now consider the implications of the baseline parameterization for optimal bank capital regulation, problem (11). According to Figure 1, bank capital regulation is effective at moving S and σ closer to the full-information regulatory optimum. For $\psi < \overline{\psi}$, D^{cap} is below the optimal choices of both the unregulated bank and the full-information regulator. In particular, we find that D^{cap} follows a "V"-shaped pattern as ψ increases. This V-shape was found in most parameterizations we have tried.

Let us consider why this result comes about. Recall that the wedge between the bank's objective and the regulator's objective is the expected value of the deposit insurance put option. This option has strike price D, so reducing D by regulatory fiat reduces the value of this option, thereby reducing the discrepancy between bank and regulatory objectives. In addition, reducing D^{cap} directly reduces the probability of default. Of course, this action is not without cost, since deposits provide liquidity value to society.

As ψ increases from zero, the value of the bank to the regulator increases; the capital regulator responds by further reducing the default probability by reducing D^{cap}. While this reduces the value of the put option, the bank still regards the put option as the main focus of value creation, since it sets σ^{cap} at $\sigma_{\max}$. However, at a franchise value $\widetilde{\psi}$ corresponding to the bottom of the "V" (around 0.1 in the baseline parameterization), D^{cap} has been reduced to the point where the put option has little value. As a result, the bank shifts discontinuously from $\sigma_{\max}$ to $\sigma_{\min}$. While the put option no longer acts as a significant wedge between bank and regulatory objectives, the bank's choices do not yet coincide with the social optimum. The distortion is now the suboptimal value of D that is imposed by the capital regulator, which in turn implies a screening level $S^{cap} < S^{reg}$, the socially optimal level. Notice that, at this point, the equilibrium default probability under optimal capital regulation is suboptimally *low.* The full-information regulator would prefer to tolerate a slightly higher default rate to avoid paying the high cost of capital implied by D^{cap}. As ψ increases from $\widetilde{\psi}$ to $\overline{\psi}$, the regulator sets D^{cap} equal to maximal value of D that induces banks to choose $\sigma_{\min}$. This value increases with ψ until, at $\overline{\psi}$, the incentives of the unregulated bank are (virtually) aligned with those of the regulator, at which point $D^{cap} \cong D^{*} \cong D^{reg}$.

While optimal capital regulation does narrow the gap between the choices of the unregulated bank and the choices of the full-information regulator, there is still room for improvement in the regulatory objective function. In section 6, below, we consider whether further improvement is possible by augmenting capital regulation with state-contingent fines.

5.2 Alternative Parameterizations

In this subsection, we perturb individual parameters away from the baseline case (leaving all other parameters at the baseline values). First, consider the effect of reducing $\mu_{\max}$ to 1.5 (implying a 50% maximal mean net return to the bank's portfolio). The behavior of this parameterization is qualitatively the same as that of the baseline parameterization. However, for low values of ψ, D^{reg} is lower. The reason is that with $\mu(S, \sigma)$ given by equation (2) the lower value of $\mu_{\max}$ reduces the marginal value of screening. As a result, the mean portfolio return is lower, so a lower value of D^{reg} is needed to control default probability. Similarly, D^{cap} is lower in this case than in the baseline case. The message here is that if a bank's investment opportunity set deteriorates, stricter capital standards are called for.

Second, suppose we double ρ to 0.1. This increases the cost of capital, so optimal D is higher. In particular, we find that $D^{reg} = D^* = 1$ for $\psi < .7$, even though $\overline{\psi}$ (the value of ψ at which σ^* shifts from $\sigma_{\max}$ to $\sigma_{\min}$) is only 0.4. Not surprisingly, D^{cap} is also higher in this case than in the baseline case.

Third, suppose we increase the marginal value of screening by setting $a_s = 10$. This increases the marginal value of screening, so S increases, implying a higher portfolio mean. This higher mean return reduces the probability of default sufficiently so that capital is no longer used by either the unregulated bank or the full-information regulator to control default. In particular, $D^* = D^{reg} = 1$ for all $\psi \in [0, 1]$. We still get the V-shaped pattern for D^{cap}, although the values of D^{cap} are higher than in the baseline case. In contrast, when we reduce a_s to 1, the marginal value of screening is so low that, for all $\psi \in [0, 1]$, $S^* = S^{reg} = S^{cap} = 0$. Now, default is controlled only by D and σ, so, unlike the baseline, $D^{reg} < 1, \forall \psi > 0$.

Fourth, let us consider the effect of positive values for a_σ. With $a_\sigma = 0.1$, both the full-information regulator and the capital regulator choose the maximal value of σ for small but positive values of ψ. That is, for low franchise value banks the benefit of increasing the portfolio mean by choosing a high value of σ exceeds the cost in increased default risk. However, we still get a reverse mean-variance trade-off in equilibrium: For the unregulated bank, the full-information regulator, and the capital regulator, σ is (weakly) decreasing in ψ while $\mu(S, \sigma)$ is strictly increasing. Since σ^{cap} now equals $\sigma_{\max}$ for low ψ (in particular, for $\psi < 0.25$ in this parameterization), the burden of reducing default risk falls primarily on capital structure. As a result, D^{cap} is lower than in the baseline parameterization. The lesson here is that if the return to greater risk increases, capital regulation must be more stringent.

When we further increase a_σ to 0.3, these effects are magnified to the point where there is a qualitative change in the implications of the model. This case is displayed in Figure 2. Now, it is optimal to take on maximal risk, *regardless* of the franchise value. That is, $\sigma^* = \sigma^{reg} = \sigma^{cap} = \sigma_{\max}$, for all ψ.[10] However, the distinction between risk-seeking and prudent banks still

[10] We tried values of ψ as high as 10,000 without finding an exception to this result.

applies. Banks with franchise value below a critical value $\overline{\psi}$ (approximately 0.51 in this parameterization) are primarily concerned with maximizing the value of the deposit insurance put option. They choose a high risk of default. At $\psi = \overline{\psi}$, default risk drops discontinuously, because the bank discontinuously increases capital from 0 to over 80%. (That is, the bank reduces D^* from 1 to 0.125.) In other words, the prudent banks control default risk by increasing capital, rather than by reducing portfolio risk. The risk-return trade-off implied by $a_\sigma = 0.3$ is so favorable that high franchise value banks act more as trading institutions than depository institutions. Similarly, the full information regulator controls risk with capital rather than by reducing σ. Notice that at $\psi = 0.07$, the optimal value of D^{reg} drops from 1 to 0.47.[11]

The final parameterization we discuss reduces the maximal and minimal variance of the bank's portfolio return. In particular, we perturb the baseline parameterization by setting $\sigma_{\max} = 0.6$ and $\sigma_{\min} = 0.1$. The main effect of lowering this risk profile is to increase optimal D (decrease capital) for the bank, full-information regulator, and capital regulator. In particular, the unregulated bank and the full-information regulator set $D = 1$ for all $\psi \in [0, 1]$. The optimal capital regulation does lower D to around 0.8 for $\psi = 0.1$, but D^{cap} returns to a value of 1 for $\psi \geq 0.2$. This result is no surprise: In an environment of lower risk, less stringent capital regulation is appropriate.

6 Results for the Optimal State-Contingent Regulator's Problem

In this section, we report results for the case where non-negative fines are allowed as an additional regulatory device. In this case, the regulator must also choose a function, $g(D, r)$, rather than simply choosing a capital requirement parameter. To obtain numerical results, we must map the infinite-dimensioned space of candidate g functions into some finite-dimensioned parameter space. One way to do so would be to assume a functional form (such as an n^{th}-order polynomial), and then to solve numerically for the optimal parameters. We found this approach to be extremely inefficient computation-

[11] This is the reason why the "V"-shaped pattern for D^{cap} no longer holds in this parameterization. D^{reg} drops to meet D^{cap} at such a low value of ψ that there is no need for D^{cap} to rise to meet D^{reg}.

ally, perhaps due to the discontinuities (documented, for example, in Figure 1) in the bank's response to small changes in the model parameters. Instead, we used versions of linear programming methods developed for mechanism design problems.[12] Specifically, we discretize both the state space and the choice sets. Following Grossman and Hart (1983), we then loop over each discrete investment and deposit choice, find the value of implementing each combinations, and then take the maximum. The problem of implementing each combination of choice variables is formulated as a linear program.[13]

The state space is discretized by imposing a set of $N + 1$ boundary points, $\{\widehat{r}_i\}_{i=1}^{N+1}$. An N-point grid over r is constructed by setting r_i equal to $E[r|\widehat{r}_i < r \leq \widehat{r}_{i+1}]$. Since the distribution of r depends on the choice variables, the state space discretization varies as the choice variables change. For example, in the baseline parameterization with $\sigma = \sigma_{\min}$, the expected return in the range of returns (2.6,20.0] was approximately 2.75, while for $\sigma = \sigma_{\max}$ it was approximately 4.23. We use this approach because it allows us to keep our discrete distribution similar to the log normal distribution. If we had kept the grid points fixed, the resulting discrete distribution would differ substantially from log normality for some values of the choice variables.

In our numerical exercises we use the following coarse grids:

- $\widehat{r} \in \{0.0001, 0.7, 0.9, 1.0, 1.1, 1.5, 2.0, 2.6, 20.0\}$
- $D \in \{0.00, 0.01, 0.02, ..., 1.00\}$
- $S \in \{0.22, 0.23, 0.24, 0.25, 0.26, 0.30\}$
- $\sigma \in \{0.3, 1.5\}$
- $\psi \in \{0.00, 0.05, 0.10, ..., 0.50\}$

Given D, the optimal fine schedule $g(D, r)$ for this problem is simply a vector with the same dimension as the grid over r. Finally, for the resource cost per unit of collected fines we set $\tau = 0.01$.

[12] See Prescott (1999) for a description of this literature and references to it.

[13] We can prove that there is no additional value to allowing randomization in the assignment of banks to deposit-investment pairs, so we do not worry about this possibility.

6.1 Results for the Basic Fines Model

In this subsection we consider the solution to problem (13) when the only restrictions fines is inequality (12). Figure 3 reports results for the baseline parameterization. It is analogous to Figure 1, except that it uses the discrete solution method and includes the solution to the optimal state-contingent regulation problem (13) as well as the solutions to problems (8), (9), and (11). There are two points to note from Figure 3. First, the optimal choices of the unregulated bank, the full-information regime, and the capital regulation regime correspond fairly closely to those displayed in Figure 1 For example, the optimal capital regulations are still V-shaped. This similarity indicates that the distortions induced by our discretization do not obscure the essential properties of our model.

Second, allowing the regulator to impose state-contingent fines in addition to capital requirements permits a significant improvement in welfare relative to the pure capital regulation case. In this example, except for $\psi = 0.15$, the optimal fines regime implements the same (discrete) values of $\{S, D, \sigma\}$ as optimal full-information regulation, so the only welfare loss would be the costs from imposing the fines. Note in particular that, unlike pure capital regulation, there is no need to impose suboptimally high capital (represented by the V-shaped configuration for D^{cap}) when the regulator can use state-contingent fines.

In Figure 4 we display the optimal fine schedule for the baseline parameterization with zero franchise value. The bars give the value of the optimal fine for each discrete value of r. Figure 4 also displays the constraint on fines imposed by limited liability (the oblique dashed line). Figure 5 is analogous to Figure 4 for the baseline parameterization with franchise value of 0.15. For both parameterizations, the $(S, \sigma) = (0.25, 0.3)$ investments were optimal and there is a large fine at the highest value of r in the discrete grid. (For this parameterization, this highest value of r is approximately 2.75, which is the mean of r, conditional on $2.6 < r \leq 20$.) In Figure 4 there is also a smaller fine at a low level of r, around 1.05. In both examples, most of the fines are imposed on the extreme right tail of the distribution. While the higher of these fines is imposed infrequently, the magnitude of the fine is quite large. With $\psi = 0$ (Figure 4) the fine approaches the limited liability constraint. That is, if this fine is triggered, it taxes away almost all of the bank's profits. With $\psi = 0.15$, the fine is lower but still substantial. According to these results, the optimal state-contingent fine mechanism is to

$(\widehat{S},\widehat{\sigma})$	[0.01,0.7]	(0.7,0.9]	(0.9,1.0]	(1.0,1.1]	(1.1,1.5]	(1.5,2.0]	(2.0,2.6]	(2.6,20.0]
(0.22,1.5)	15,808.0	44.2	5.08	1.62	0.35	0.26	1.09	52.16
(0.24,0.3)	1.31	1.18	1.13	1.09	1.02	0.98	0.96	0.95

Table 1: Selected likelihood ratios when $(S,\sigma) = (0.25, 0.3)$ is implemented.

impose large fines at the extreme upper tail.

One might ask why a fine is effective if it is imposed so infrequently. In particular, the upper-tail fine in Figure 4 is only imposed when $r \in (2.6, 20]$, an event that happens with probability 0.0026 in equilibrium (where the bank's choices are $(S,\sigma) = (0.25, 0.3)$). This fine is effective in altering bank incentives because the probability of being fined would be much higher if the bank were to make the suboptimal choice $\sigma = \sigma_{\max}$. Specifically, if the bank chose $(S,\sigma) = (0.22, 1.5)$ the probability of this return would be 0.1442. In other words, a fine is most effective the less frequently it affects a bank that is doing what it is supposed to do and the more frequently it affects a bank that is not doing what it is supposed to do.

More formally, we can understand why fines are assessed at the right-hand tail if we consider the *likelihood ratio,* the ratio of the probability under an alternative choice to that under the recommended choice. The importance of likelihood ratios in moral hazard models is well known.[14] Suppose (S,σ) is the investment decision the regulator wants to implement and $(\widehat{S},\widehat{\sigma})$ is the investment decision the regulator wants to prevent. The most cost-effective incentive mechanism is to punish the agent is in those states where the likelihood ratio $\frac{f(r|\widehat{S},\widehat{\sigma})}{f(r|S,\sigma)}$ is high. At these values of r, the fine function punishes deviating banks much more frequently than banks who conform to the regulator's wishes. That is, there is the most prevention per dollar of fine assessed in equilibrium.

Table 1 shows the likelihood ratios for two possible investment deviations $(\widehat{S},\widehat{\sigma})$ when $(S,\sigma) = (0.25, 0.3)$ is implemented. The first row lists the likelihood ratios if the low-screening, high-variance investment, $(\widehat{S},\widehat{\sigma}) = (0.22, 1.5)$, is taken. The second row lists the ratios for a low-screening, low-variance investment. In the first row, the likelihood ratio is actually highest on the left-hand side of return distribution, but because of limited liability no fines can be assessed there. Instead, we see the fines imposed at the

[14]See, for example, Hart and Holmstrom (1987).

right-hand sided of the distribution.

We still have not explained why the regulator wants to prevent the (0.22,1.5) investment rather than, say, a higher-screening, high-variance investment. For banks with low franchise value, the regulator does not want to prevent risk per se, but it does want to prevent low screening. The reason the low-screening, high-variance investment choice is such a concern to the regulator is that deposit insurance gives banks a preference for risk. If the bank takes a high-risk investment, however, there is a higher probability of insolvency and the marginal benefit to the bank of screening declines. As a result, banks that choose high σ also tend to choose low values of S.

Why is there a fine assessed at the lowest feasible return for the 0.0 franchise value case, described in Figure 4? The reason for this fine is extremely interesting and demonstrates how regulations designed to prevent one potential problem may cause another problem along some other dimension. As we just discussed, the fines on the right-hand tail are designed to stop the bank from taking the low-screening, high-variance investment. But this deviation is not the only incentive constraint that binds. The $(S, \sigma) = (0.24, 0.3)$ incentive constraint binds as well. As row 2 in Table 1 demonstrates, the most effective feasible return in which the regulator can impose a fine to prevent this investment is the (1.0,1.1] return. The regulator would like to levy fines at lower levels of the return but, again, limited liability makes that infeasible.

The reason why a fine is needed to prevent suboptimally low screening is not as apparent as it may seem. At low variance levels, the marginal trade-offs of screening for the bank and the regulator are very closely aligned. We ran an experiment for the baseline parameterization where we fixed the variance at 0.3, and let banks choose only the screening level. In this experiment, the screening level of 0.25 is implemented *without* any fines being imposed. This may be surprising since the incentive constraint on the $(\widehat{S}, \widehat{\sigma}) = (0.24, 0.3)$ investment choice was binding in the baseline case. What is going is very subtle. As we said earlier, for low variance investments the bank and regulator's incentives are roughly aligned along the screening dimension. When we add variance as a choice variable, the regulator imposes high-return fines to prevent the disastrous low-screening, high-variance investment. But in this example, once a fine is assessed it also affects the bank's marginal incentive to screen at the low level of variance! Even though the fine is imposed infrequently, it is large enough in magnitude to reduce the marginal benefit from screening, leading banks to prefer a suboptimally low value of S. As a result, the bank's and regulator's marginal incentives are no longer as closely aligned. By adding fines to prevent the low-screening, high-variance investment, the regulator has created an incentive for the bank to take a low-variance, low-screening investment. Consequently, an additional fine at the low return level is also needed.

To summarize, there are two noteworthy features of this example. First, the combination of variance and screening drives the incentive problem. This is particularly evident at a franchise value of zero where there is no incentive problem along the variance dimension. Second, the regulatory response of a high-return fine to prevent the low-screening, high-variance deviation creates a need for a low-return fine in order to prevent a deviation along the screening low-variance dimension, which was not otherwise an incentive concern.

The combination of low screening and high variance is the main regulatory concern in this problem because with high risk investments, the probability of insolvency goes up, diminishing the value of screening to the bank. As the franchise value increases, we would expect this feature to become less of a concern. This is precisely what we found. As franchise value increases, the size of the fines imposed on the right-hand side of the tail monotonically decreases. Initially we see the fines at the lower tail return disappear until the franchise value reaches 0.20, at which point the next highest level of screening is implemented and the low-level fines return. Again, however, these fines disappear as the franchise value increases and then all fines disappear by the time the franchise value reaches 0.40.

6.2 Practical Considerations in Implementing State-Contingent Fines

Do the results of this section offer any guidance to policy makers? It may seem politically infeasible to fine banks, although fines were considered under the Federal Reserve Board's pre-commitment proposal. Fines at the lower tail of the distribution do resemble the punitive features of the prompt corrective action provisions of FDICIA. It might be difficult, however, to impose fines when a bank performs exceptionally well. The bank could argue that such a fine schedule penalizes skillful management rather than excessive risk. Still, fines do resemble warrants which are sometimes issued in conjuction with private debt contracts. A warrant gives the holder the option to purchase stock at a fixed price. If the warrant price is set high enough, then holders will only exercise them when the issuing firm does extremely well, thus penalizing, in a limited sense, firm risk taking. Regardless of whether or not fines are feasible in practice, at a minimum these findings suggest that extremely high returns should act as a red flag to regulators. Instead of a fine, regulators could use the results to decide which banks to target their supervisory resources. For example, a highly detailed audit could be conducted and if ex ante high risk strategies were ascertained then regulatory sanctions could be imposed.[15]

A more serious concern with the results in our examples is that the fines are narrowly targeted at returns with high likelihood ratios. In practice, banks have some control over their reported earnings. Large targeted fines would give banks incentives to alter reported numbers, even to raise expenses so as to lower returns out of the fine range. To handle this concern would require extending the model so that banks have some control over reported returns. One simple way to do this is to allow banks to costlessly destroy any portion of their return that they want. Of course, banks would only do so if the net return, $r - g(D, r)$, is decreasing in r over some realizations of the return. Therefore, the optimal state-contingent regulatory problem in this environment is equivalent to problem (13) with the additional constraint that $r - g(D, r)$ be monotonically increasing in the return r.

Figure 6 summarizes the results for this problem under the baseline parameterization with this monotonicity constraint imposed. For purposes of comparison, Figure 6 also includes the results from Figure 3 for the capital regulation and the capital regulation with fines cases. According to Figure 6, limiting net returns to be monotonic reduces the efficacy of fines but the effect is only dramatic at a zero franchise value. Note in particular that the regulator can no longer induce the zero franchise value bank to choose $\sigma = \sigma_{\min}$ when monotonicity is imposed. While capital ratios are close to those of the capital and fines case, the monotonicity requirement does imply more stringent capital requirements at $\psi = 0.05$. This suggests that the more fines are limited the closer the problem will be to the pure capital regulation case.

Interestingly, despite the superiority of the monotonicity constraint case to pure capital regulations, we find that for low franchise values, screening is higher under capital regulations. Apparently, the fines are effective at preventing the low-screening, high-variance deviation but imposing these fines

[15] Conversations with regulatory personnel indicate that they do this implicitly. For example, fast growth is considered a potential flag for safety and soundness concerns.

makes the low-screening, low-variance investment attractive enough that the banks make that choice.

Figure 7 reports optimal fine schedules and net returns for the monotonicity case under the baseline parameterization with $\psi = 0.05$. The top panel lists the fine schedule and the bottom panel lists the bank's net return. In the net returns panel, we leave out franchise value and report, $r - g(D, r) - D$. In this figure, we can see the effect of the monotonicity constraint. The program wants to impose fines at the highest range of returns, where the likelihood ratio is high, but it is limited by the monotonicity constraint. Indeed, without the monotonicity constraint all fines would be imposed at the highest return for this parameterization. However, with the monotonicity constraint, this strategy is not feasible, so it also imposes fines at the second highest level of output. But now, these fines affect the desirability of low-screening, low-variance investments and as we saw earlier, the problem must also impose fines at the lower end of solvency range to prevent these deviations.

Costless destruction of returns is not the only way one can model the idea that bank's have some control over their reported returns. Another approach is to follow Lacker and Weinberg (1989) and allow the bank to change the observed return at a cost. Ultimately, the more leeway the bank has in reporting its returns the less effective fines will be and the more reliant regulators will be on ex ante regulations like capital requirements. In general, the more control the bank has over final returns, be it through control over reported returns or through control of a more general family of return distributions,[16] the less effective fines will be. Still, even in this modern era of financial derivatives, there are some limits on the ability of institutions to control final returns so further research into these limitations and the effectiveness of state-contingent penalties seems warranted.

7 Conclusions and Directions for Future Research

In this paper, we find that the put option associated with deposit insurance induces low franchise value banks to choose one or more of the following so-

[16]The most extreme case would be if there existed a complete set of state-contingent claims, in which case the bank could sell forward all earnings (above the solvency constraint) in those states where fines were assessed. In this sense, our results clearly require a degree of market incompleteness.

cially inefficient actions: suboptimally low capital; suboptimally high portfolio risk; and/or suboptimally low screening effort. These inefficient choices are socially harmful because they imply a suboptimally high default probability and a suboptimally low mean return to bank assets. We find that capital requirements can be useful in reducing these distortions. Capital requirements directly reduce the default probability, and also act to induce banks to increase screening activity and to reduce portfolio risk. However, capital requirements are not without social cost, so capital requirements alone generally do not recover the first-best allocation. When augmented with state-contingent penalties, capital regulation comes much closer to the first-best. In particular, there is not always a need for suboptimally high capital. However, the particular form taken by the optimal fine schedule is somewhat unusual: It involves very high fines at the extreme upper tail of the distribution, often accompanied by small fines just above the bankruptcy point. Fines of this type are, to our knowledge, not used in practice, though as we discussed earlier, they resemble warrants and can be replicated with other ex post regulatory devices, such as in-depth costly audits.

Two extensions of this model seem to be of particular interest. First, our model does not consider general equilibrium effects of regulation. In particular, the return to deposits is fixed at unity, and we do not require banks to sell equity capital to households. Equilibrium effects of regulation are of interest because the price of equity capital would depend on the expected payoff to the bank, which is clearly endogenous to the regulatory regime. A second extension would be to relax the assumption that bank characteristics are observable to the regulators. It would be of interest to assume that both bank characteristics, such as franchise value, and bank actions are unobservable. For example, capital requirements that decrease in franchise value, as we saw for a range of franchise values in Figure 3, might be difficult to support. It would also be interesting to allow banks' technologies to differ in such a way that some banks were more productive than others. In that kind of environment, high return fines might be less appealing since it would be hard to distinguish between a high productivity bank and an imprudent bank on the basis of high returns. Little work has been done to address these sorts of questions in a bank regulation model. Our numerical methods are potentially suitable to conduct this analysis.

A Proofs of Propositions

Let $F^{\ln}(\cdot|\mu,\sigma)$ and $f^{\ln}(\cdot|\mu,\sigma)$ denote respectively the log normal cdf and pdf with mean μ and variance σ^2. The following Lemma is used in the proofs of the propositions.

Lemma 4 *If* $0 < D < \mu$, *then*

$$F_{\mu}^{\ln}(D|\mu,\sigma) < 0, \text{ and}$$

$$F_{\sigma}^{\ln}(D|\mu,\sigma) > 0.$$

(Part 1 of proposition 1, along with the constraint $D \leq 1$ imply that $0 < D^* < \mu_{\min}$ as long a all portfolios have positive expected net return (a natural specification). Therefore, the conditions on D required for lemma 4 hold at the bank's optimal choices.)

Proof of Lemma 4:

$$F^{\ln}(D|\mu,\sigma) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{1}{\sqrt{2}}\frac{\log(D) - \mu_n}{\sigma_n}\right)\right] \tag{14}$$

where the error function erf is defined by

$$\operatorname{erf}(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt,$$

$$\mu_n \equiv \frac{1}{2}\log\left(\frac{\mu^4}{\mu^2+\sigma^2}\right), \qquad \sigma_n \equiv \sqrt{\log\left(\frac{\mu^2+\sigma^2}{\mu^2}\right)}.$$

For a random variable x with cdf $F^{\ln}(x|\mu,\sigma)$, μ_n and σ_n give the mean and standard deviation of $\log(x)$, respectively. Differentiating equation (14) with respect to μ and σ, one obtains

$$F_{\mu}^{\ln}(D|\mu,\sigma) = -Q(D,\mu,\sigma)\frac{\sigma}{\mu}\left(2 + \frac{\mu^2}{\sigma^2} - \left(\frac{\log(D) - \mu_n}{\sigma_n^2}\right)\right)$$

$$F_{\sigma}^{\ln}(D|\mu,\sigma) = Q(D,\mu,\sigma)\left(1 - \left(\frac{\log(D) - \mu_n}{\sigma_n^2}\right)\right)$$

where

$$Q(D,\mu,\sigma) \equiv \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\log(D)-\mu_n}{\sigma_n}\right)^2}\left(\frac{\sigma}{\sigma_n(\mu^2+\sigma^2)}\right) > 0. \tag{15}$$

In equation (15), the strict positivity of Q follows from $D > 0$ To prove the lemma, it is sufficient to show that

$$\left(2+\frac{\mu^2}{\sigma^2}-\left(\frac{\log(D)-\mu_n}{\sigma_n^2}\right)\right)>0 \tag{16}$$

and

$$\left(1-\left(\frac{\log(D)-\mu_n}{\sigma_n^2}\right)\right)>0 \tag{17}$$

We first demonstrate equation (16).

$$sign\left[2+\frac{\mu^2}{\sigma^2}-\left(\frac{\log(D)-\mu_n}{\sigma_n^2}\right)\right]=sign\left[\sigma_n^2-\log(D)+\mu_n+\sigma_n^2\left(\frac{\mu^2}{\sigma^2}+1\right)\right].$$

The term $\sigma_n^2\left(\frac{\mu^2}{\sigma^2}+1\right)$ is positive, so to prove (16) it is sufficient to show that $\sigma_n^2-\log(D)+\mu_n>0$ as follows:

$$\begin{aligned}\sigma_n^2-\log(D)+\mu_n &= \log\left(\frac{\mu^2+\sigma^2}{\mu^2}\right)-\log(D)+\frac{1}{2}\log\left(\frac{\mu^4}{\mu^2+\sigma^2}\right)\\ &= \log\left(\frac{\sqrt{\mu^2+\sigma^2}}{D}\right)>0 \end{aligned} \tag{18}$$

since $D<\mu$. Similarly, to demonstrate equation (17), note that

$$sign\left[1-\left(\frac{\log(D)-\mu_n}{\sigma_n^2}\right)\right]=sign\left[\sigma_n^2-\log(D)+\mu_n\right]>0$$

by equation (18). This completes the proof of the Lemma.

A.1 Proof of Proposition 1

The first-order and second-order necessary conditions for the solution to problem (8) are as follows:

First order condition with respect to D :

$$\rho+F\left(D|S,\sigma\right)-\psi f\left(D|S,\sigma\right)\left\{\begin{array}{l}\leq 0 \text{ if } D=0\\ =0 \text{ if } 0<D<1\\ \geq 0 \text{ if } D=1\end{array}\right\} \tag{19}$$

If D is interior, the second order necessary condition is

$$f\left(D|S,\sigma\right)-\psi f'\left(D|S,\sigma\right)<0. \tag{20}$$

First order condition with respect to σ :

$$\int_0^D \left(D-r-\psi\right) f_\sigma\left(r|S,\sigma\right) dr+\mu_\sigma \left\{\begin{array}{l} \leq 0 \text{ if } \sigma=\sigma_{\min} \\ =0 \text{ if } \sigma_{\min}<\sigma<\sigma_{\max} \\ \geq 0 \text{ if } \sigma=\sigma_{\max} \end{array}\right\} \tag{21}$$

If σ is interior, the second order necessary condition is

$$\int_0^D \left(D-r-\psi\right) f_{\sigma\sigma}\left(r|S,\sigma\right) dr+\mu_{\sigma\sigma}<0 \tag{22}$$

First order condition with respect to S :

$$\int_0^D \left(D-r-\psi\right) f_S\left(r|S,\sigma\right) dr+\mu_S-\gamma \left\{\begin{array}{l} \leq 0 \text{ if } S=0 \\ =0 \text{ if } S>0 \end{array}\right\} \tag{23}$$

If σ is interior, the second order necessary condition is

$$\int_0^D \left(D-r-\psi\right) f_{SS}\left(r|S,\sigma\right) dr+\mu_{SS}<0. \tag{24}$$

We now use these first- and second-order conditions to prove proposition 1.

Proposition 1.1:

$F(0|S,\sigma)=f(0|S,\sigma)=0$, but $\rho>0$, so at $D=0$ the left-hand side of equation (19) is strictly positive.

Proposition 1.2:

If $\psi=0$, the left-hand side of equation (19) is strictly positive for all D.
Proposition 1.3:

If $D^*$is interior, then equation (19) implies

$$\rho+F\left(D^*|S^*,\sigma^*\right)-\psi f\left(D^*|S^*,\sigma^*\right)\equiv 0. \tag{25}$$

To prove proposition 1.3, differentiate equation (25) with respect to ψ and ρ, and impose inequality (20).

Proposition 1.4:

Rewrite the left-hand side of equation (23) as

$$\mu_S \left[\int_0^D (D - r - \psi) f_\mu^{\ln} (r|\mu, \sigma)\, dr + 1 \right] - \gamma$$

The proposition follows from $\left[\int_0^D (D - r - \psi) f_\mu^{\ln} (r|\mu, \sigma)\, dr + 1 \right] < \infty$.

Proposition 1.5:

If S^* is interior, then equation (23) implies

$$\int_0^{D^*} (D^* - r - \psi) f_S (r|S^*, \sigma^*)\, dr + \mu_S (S^*, \sigma^*) \equiv 0. \qquad (26)$$

To prove proposition 1.5, differentiate equation (26) with respect to ψ and use $F_S(D|S, \sigma) = F_\mu^{\ln}(D|\mu, \sigma)\mu_S$ to get

$$\frac{\partial S^*}{\partial \psi} = \frac{F_\mu^{\ln}(D|\mu, \sigma)\mu_S}{\int_0^D (D - r - \psi) f_{SS} (r|S, \sigma)\, dr + \mu_{SS}} > 0 \qquad (27)$$

where the last inequality in equation (27) follows from equation (24), lemma 4, and $\mu_S > 0$. The proof that $\frac{\partial S^*}{\partial \gamma} < 0$ is analogous.

This ends the proof.

A.2 Proof of Proposition 2

Marshall and Venkataraman (1999) proved proposition 2 for the case where $\mu_\sigma = 0$. The generalization to μ_σ satisfying equation (33) is straightforward, and is sketched.

Proof of proposition 2.1:

It is sufficient to show that if equation (21) holds at equality then inequality (22) is violated. If equation (21) held at equality, then

$$D - \psi = \frac{\int_0^D r f_\sigma (r|S, \sigma)\, dr - \mu_\sigma}{F_\sigma (D|S, \sigma)} \qquad (28)$$

Substitute equation (28) into the left-hand side of inequality (22) and rearranging, one obtains

$$\left[F_{\sigma\sigma}\left(D|S,\sigma\right)\int_0^D rf_\sigma\left(r|S,\sigma\right)dr - F_\sigma\left(D|S,\sigma\right)\int_0^D rf_{\sigma\sigma}\left(r|S,\sigma\right)dr\right]$$
$$+\left[\mu_{\sigma\sigma}F_\sigma\left(D|S,\sigma\right) - \mu_\sigma F_{\sigma\sigma}\left(D|S,\sigma\right)\right] \quad (29)$$

To prove proposition 2.1, it is sufficient to show that for sufficiently small μ_σ and $\mu_{\sigma\sigma}$ the object in (29) is strictly positive, implying that inequality (22) is violated. To that end, note that

$$\lim_{\mu_\sigma,\mu_{\sigma\sigma}\to 0}\left[\mu_{\sigma\sigma}F_\sigma\left(D|S,\sigma\right) - \mu_\sigma F_{\sigma\sigma}\left(D|S,\sigma\right)\right] = 0 \quad (30)$$

and

$$\lim_{\mu_\sigma,\mu_{\sigma\sigma}\to 0}\left[F_{\sigma\sigma}\left(D|S,\sigma\right)\int_0^D rf_\sigma\left(r|S,\sigma\right)dr - F_\sigma\left(D|S,\sigma\right)\int_0^D rf_{\sigma\sigma}\left(r|S,\sigma\right)dr\right]$$

$$= F_{\sigma\sigma}^{\ln}\left(D|\mu,\sigma\right)\int_0^D rf_\sigma^{\ln}\left(r|\mu,\sigma\right)dr - F_\sigma^{\ln}\left(D|\mu,\sigma\right)\int_0^D rf_{\sigma\sigma}^{\ln}\left(r|\mu,\sigma\right)dr \quad (31)$$

Finally,

$$F_{\sigma\sigma}^{\ln}\left(D|\mu,\sigma\right)\int_0^D rf_\sigma^{\ln}\left(r|\mu,\sigma\right)dr - F_\sigma^{\ln}\left(D|\mu,\sigma\right)\int_0^D rf_{\sigma\sigma}^{\ln}\left(r|\mu,\sigma\right)dr > 0$$

if and only if

$$\frac{\partial}{\partial\sigma}\left[\frac{\int_0^D rf_\sigma^{\ln}\left(r|\mu,\sigma\right)dr}{F_\sigma^{\ln}\left(D|\mu,\sigma\right)}\right] < 0. \quad (32)$$

Inequality (32) is proved in Marshall and Venkataraman (1999, Lemma A.2, p. 152). Together, equations (30) - (32) imply that the object in equation (29) is strictly positive as long as both μ_σ and $\mu_{\sigma\sigma}$ are sufficiently small.

Proof of proposition 2.2:

We first note that if μ_σ is small enough, $F_\sigma(D|S,\sigma)$ is unambiguously positive. In particular, suppose

$$\max\left[\mu_\sigma\right] < \min\left[-\frac{F_\sigma^{\ln}(D|\mu,\sigma)}{F_\mu^{\ln}(D|\mu,\sigma)}\right] \quad (33)$$

where the minimization in (33) is taken over $D \in [0,1]$, $\mu \in [\mu_{\min}, \mu_{\max}]$, $\sigma \in [\sigma_{\min}, \sigma_{\max}]$.[17] If inequality (33) holds, then

$$F_\sigma(D|S,\sigma) = F_\sigma^{\ln}(D|\mu,\sigma) + \mu_\sigma F_\mu^{\ln}(D|\mu,\sigma) > 0 \tag{34}$$

where the inequality in (34) follows from Lemma 4 in Appendix A.

Using integration by parts, the left-hand side of equation (21) can be written

$$\int_0^D F_\sigma\left(D|S,\sigma\right) dr - \psi F_\sigma\left(D|S,\sigma\right). \tag{35}$$

If $\psi = 0$ and inequality (33) holds, then lemma 4 and proposition 1.1 imply that the object in (35) is strictly positive.

Proof of proposition 2.3:

Let

$$\overline{\psi} = \frac{\int_0^D F_\sigma\left(D|S,\sigma_{\max}\right) dr}{F_\sigma\left(D|S,\sigma_{\max}\right)}.$$

Lemma A.3 in Marshall and Venkataraman (1999) shows that $\frac{\int_0^D F_\sigma^{\ln}(D|\mu,\sigma) dr}{F_\sigma^{\ln}(D|\mu,\sigma)} < \infty$. Therefore, $\overline{\psi} < \infty$ for small enough μ_σ. Use representation (35) for the left-hand side of first-order condition (21). If $\psi < \overline{\psi}$, the first-order condition implies that $\sigma^* = \sigma_{\max}$. If $\psi > \overline{\psi}$, representation (35) implies that the left-hand side of (21) is strictly negative when evaluated at $\sigma_{\max}$, implying that $\sigma^* \neq \sigma_{\max}$. Proposition 2.1 then implies that $\sigma^* = \sigma_{\min}$.

This ends the proof.

A.3 Proof of Proposition 3

The first- and second-order necessary conditions for a solution to problem (9) are as follows:

[17] The right-hand side of equation (33) is strictly greater than zero, even though $F_\sigma^{\ln}(0|\mu,\sigma) = F_\mu^{\ln}(0|\mu,\sigma) = 0$. Using L'Hôpital's rule, one can show that

$$\lim_{D \to 0} \left| \frac{F_\sigma^{\ln}(D|\mu,\sigma)}{F_\mu^{\ln}(D|\mu,\sigma)} \right| = \frac{\sigma}{\mu} > 0.$$

First order condition with respect to D :

$$\rho - \psi f\left(D|S,\sigma\right)\left\{\begin{array}{l} \leq 0 \text{ if } D=0 \\ = 0 \text{ if } 0<D<1 \\ \geq 0 \text{ if } D=1 \end{array}\right\} \tag{36}$$

If D is interior, the second order necessary condition is

$$-\psi f'\left(D|S,\sigma\right) < 0. \tag{37}$$

First order condition with respect to σ :

$$\mu_\sigma - \psi F_\sigma(D|S,\sigma)\left\{\begin{array}{l} \leq 0 \text{ if } \sigma=\sigma_{\min} \\ = 0 \text{ if } \sigma_{\min}<\sigma<\sigma_{\max} \\ \geq 0 \text{ if } \sigma=\sigma_{\max} \end{array}\right\} \tag{38}$$

If σ is interior, the second order necessary condition is

$$\mu_{\sigma\sigma} - \psi F_{\sigma\sigma}(D|S,\sigma) < 0 \tag{39}$$

First order condition with respect to S :

$$\mu_S - \gamma - \psi F_S\left(r|S,\sigma\right)\left\{\begin{array}{l} \leq 0 \text{ if } S=0 \\ = 0 \text{ if } S>0 \end{array}\right\} \tag{40}$$

If S is interior, the second order necessary condition is

$$\mu_{SS} - \psi F_{SS}(D|S,\sigma) < 0 \tag{41}$$

We now prove proposition 3

Proposition 3.1:

Evaluate the left-hand side of equation (36) at $D = 0$.

Proposition 3, parts 2,3, and 4:

These results follow by differentiating the relevant first-order condition at equality and using the relevant second-order condition. Also used are Lemma 4, and the facts that $F_\sigma(D|s,\sigma) > 0$ if μ_σ satisfies (33) and that

$$F_s(D|s,\sigma) = \mu_s F_\mu^{\ln}(D|\mu,\sigma).$$

This ends the proof.

References

[1] Benston, George J. and Mike Carhill. "The Causes of Consequences of the Thrift Disaster." In *Research in Financial Services: Private and Public Policy* (Ed. George G. Kaufman), JAI Press Inc., Greenwich, 1994,103-168.

[2] Berger, Allen N. "The Relationship between Capital and Earnings in Banking," *Journal of Money, Credit, and Banking,* 27, May 1995, pp. 432-456.

[3] Berger, Allen N., Herring, Richard J., and Giorgio P. Szego. "The Role of Capital in Financial Institutions," *Journal of Banking and Finance*, 19, (1995), pp. 393-430.

[4] Besanko, David, and George Kanatas. "The Regulation of Bank Capital: Do Bank Capital Standards Promote Bank Safety," *Journal of Financial Intermediation*, 5, April 1996, pp. 160-183.

[5] Bhattacharya, Sudipto, Arnoud W.A. Boot, and Anjan V. Thakor. "The Economics of Bank Regulation." *Journal of Money, Credit, and Banking* 30, November 1998, pp. 745-770.

[6] Boyd, John H., Chun Chang, and Bruce D. Smith. "Moral Hazard under Commercial and Universal Banking." *Journal of Money, Credit, and Banking* 30, August 1998, part 2, pp. 426-471.

[7] Boyd, John H., Chun Chang, and Bruce D. Smith. "Deposit Insurance: A Reconsideration." FRB Minneapolis, Working Paper, No. 593, December 1998.

[8] Boyd, John H., and Mark Gertler. "The Role of Large Banks in the Recent U.S. Banking Crisis." *Federal Reserve Bank of Minneapolis Quarterly Review*, 18, (Winter 1994).

[9] Calem, Paul and Rafeal Rob. "The Impact of Capital-Based Regulation on Bank Risk-Taking." *Journal of Financial Intermediation*, 8, 1999, 317-352. 1999.

[10] Campbell, Tim S., Yuk-Shee Chan, and Anthony M. Marino. "An Incentive-Based Theory of Bank Regulation." *Journal of Financial Intermediation*, 2, 1992, pp. 255-276.

[11] Chan, Y.S., Stuart Greenbaum, and Anjan Thakor. "Is Fairly Priced Deposit Insurance Possible?" *Journal of Finance* 47, 1992, pp. 51-60.

[12] Demsetz, Rebecca S., and Marc R. Saidenberg, and Philip E. Strahan. "Banks with Something to Lose: The Disciplinary Role of Franchise Value," Federal Reserve Bank of New York *Economic Policy Review*, 2, 1996, pp. 1-14.

[13] Dewatripont, M. and Jean Tirole. *The Prudential Regulation of Banks*, MIT Press, Cambridge, MA, 1993.

[14] Diamond, Douglas, and Philip Dybvig. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy,* 91, 1983, pp. 401-419.

[15] Diamond, Douglas and Raghuram Rajan. "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking", mimeo, University of Chicago, 1998.

[16] Freixas, Xavier, and Jean-Charles Rochet. *Microeconomics of Banking*, MIT Press, Cambridge, MA, 1997.

[17] Gan, Jie,. "The Capital-Structure and Asset-Risk Decisions of Financial Institutions: A Theory and Some Evidence." Sloan School of Management, MIT, Manuscript, 1999.

[18] Giammarino, Ronald M., Tracy R. Lewis, and David E.M. Sappington. "An Incentive Approach to Banking Regulation," *Journal of Finance* 48, September 1993, pp. 1523-1542.

[19] Gorton, Gary and Andrew Winton. "Liquidity Provision and the Social Cost of Bank Capital," Carlson School of Management, University of Minnesota, Working Paper, September 1998.

[20] Grossman, Sanford J. and Oliver D. Hart. "An Analysis of the Principal-Agent Problem." *Econometrica* 51 (January 1983): 7-45.

[21] Hart, Oliver and Bengt Holmstrom. "The Theory of Contracts." In *Advances in Economic Theory: Fifth World Congress.* (Ed. Truman F. Bewley) Cambridge, England: Cambridge University Press, 1987, pp. 71-155.

[22] Jones, David. "Emerging Problems with the Basel Capital Accord: Regulatory Capital Arbitrage and Related Issues", J*ournal of Banking & Finance*, 24, 2000, 35-38.

[23] Kareken, John H. and Neil Wallace. "Deposit Insurance and Bank Regulation: A Partial-Equilibrium Exposition." *Journal of Business,* 51, July 1978, pp. 413-438.

[24] Keeley, Michael C. "Deposit Insurance, Risk, and Market Power in Banking." *American Economic Review*, 80, December 1990, pp. 1183-1200.

[25] Kupiec, Paul H. and James O'Brien. "Recent Developments in Bank Capital Regulation of Market Risks," Working Paper 95-51. Washington: Board of Governors of the Federal Reserve System, December 1995a.

[26] Kupiec, Paul H. and James O'Brien. "A Pre-commitment Approach to Capital Requirements for Market Risk," Working Paper 95-36. Washington: Board of Governors of the Federal Reserve System, July 1995b.

[27] Lacker, Jeffrey and John A. Weinberg. "Optimal Contracts under Costly State Falsification." *Journal of Political Economy*, 97, December 1989, pp. 1345-1363.

[28] Marshall, David and Subu Venkataraman. "Bank Capital Standards for Market Risk: A Welfare Analysis," *European Finance Review,* 2, 1998, pp. 125-157.

[29] Matutes, Carmen, and Xavier Vives. "Imperfect Competition, Risk Taking, and Regulation in Banking," *European Economic Review*, 44, 2000, pp. 1-34.

[30] Mirrlees, James. "The Theory of Moral Hazard and Unobservable Behavior: Part I", *Review of Economic Studies*, 66, 1999, 3-21.

[31] Merton, Robert. "An Analytic Derivation of the Cost of Deposit Insurance Guarantees." *Journal of Banking and Finance*, 1, 1977, pp. 3-11.

[32] Mingo, John J. "Policy Implications of the Federal Reserve Study of Credit Risk Models at Major US Banking Institutions", *Journal of Banking and Finance*, 24, 2000, 15-33.

[33] Nagarajan, S. and C.W. Sealey. "State-Contingent Regulatory Mechanisms and Fairly Priced Deposit Insurance," *Journal of Banking and Finance* 22, 1998, pp. 1139-1156.

[34] Prescott, Edward S. "The Precommitment Approach in a Model of Regulatory Bank Capital." Federal Reserve Bank of Richmond *Economic Quarterly*, 83, Winter 1997, pp. 23-50.

[35] Prescott, Edward S. "A Primer on Moral Hazard Models." Federal Reserve Bank of Richmond *Economic Quarterly*, 85, Winter 1999, 47-77.

[36] Rochet, Jean-Charles. "Capital Requirements and the Behavior of Commercial Banks," *European Economic Review* 36, 1992, pp. 1137-1178.

[37] Rochet, Jean-Charles. "Solvency Regulations and the Management of Banking Risks," *European Economic Review* 43, 1999, pp. 981-990.

[38] Treacy, William F. and Mark Carey. "Credit Risk Rating Systems at Large US Banks." *Journal of Banking and Finance*, 24, 2000, pp. 167-201.

[39] White, Lawrence J. *The S&L Debacle*, Oxford University Press, New York, 1991.

[40] Williamson, Stephen D. "Discount Window Lending and Deposit Insurance," *Review of Economic Dynamics* 1 (1998), pp. 246-275.

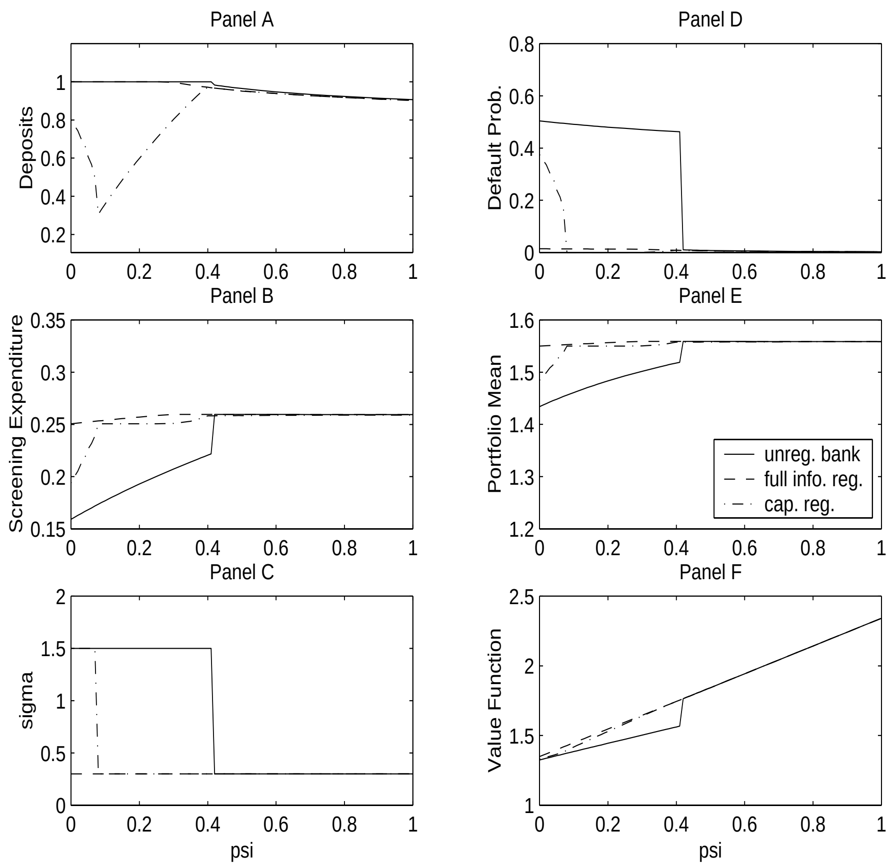

Figure 1: Baseline parameterization: $rf = 1.05$; $\rho = 0.05$; $\mu_{\max} = 1.75$; $\mu_{\min} = 1.05$; $\sigma_{\max} = 1.5$; $\sigma_{\min} = 0.3$; $a_s = 5$; $a_\sigma = 0$. For franchise value ψ ranging from zero to 1 (horizontal axes), this figure displays the deposit levels (Panel A), screening expenditure (Panel B), portfolio risk σ (Panel C), default probability (Panel D), portfolio mean (Panel E), and value of regulator's objective function (Panel F) implied by the following three regulatory regimes: Unregulated bank (solid line); optimal full-information regulation (dashed line); optimal capital regulation (dash-dot line).

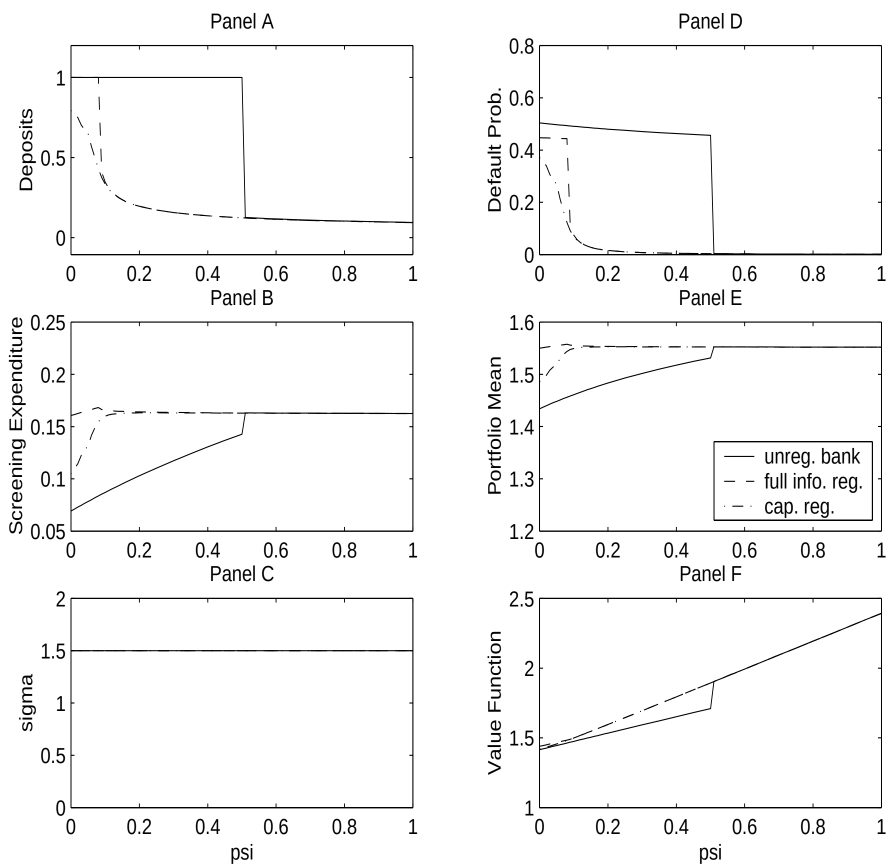

Figure 2: Baseline parameterization with positive mean-variance tradeoff. This figure modifies the baseline parameterization in Figure 1 by setting $a_\sigma = 0.3$.

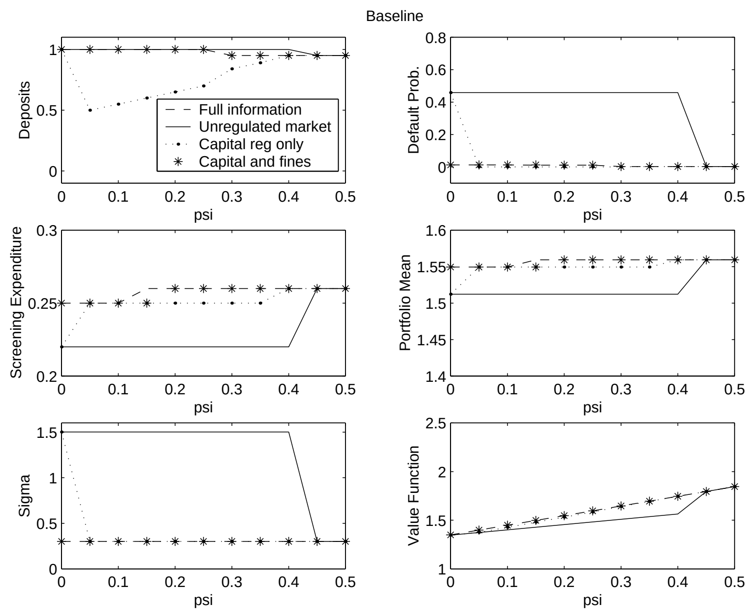

Figure 3: Summary statistics for baseline run using gridded methods. Includes solution to the optimal state-contingent problem (represented by the stars).

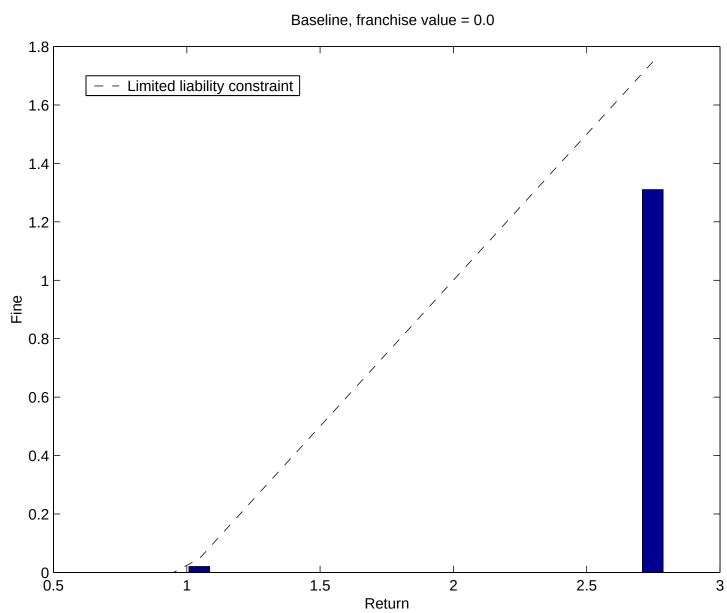

Figure 4: Optimal fine schedule to the state-contingent problem for the baseline run with $\psi = 0.0$.

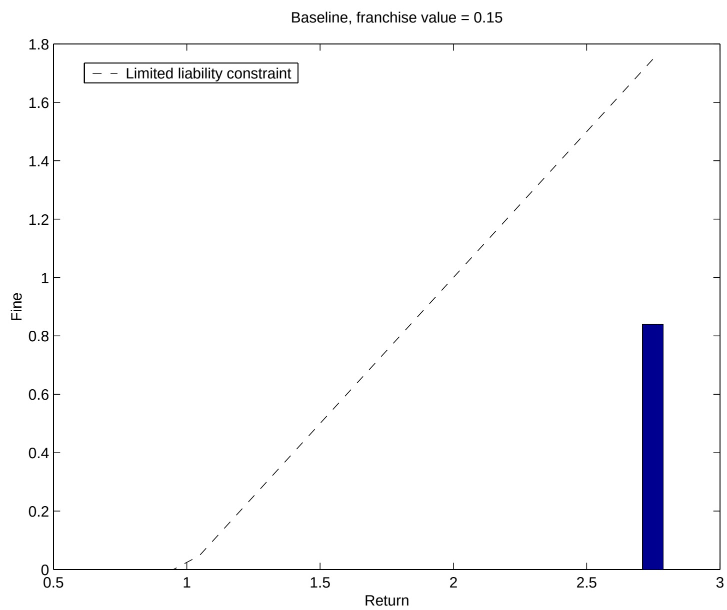

Figure 5: Optimal fine schedule to the state-contingent problem for the baseline run with $\psi = 0.15$.

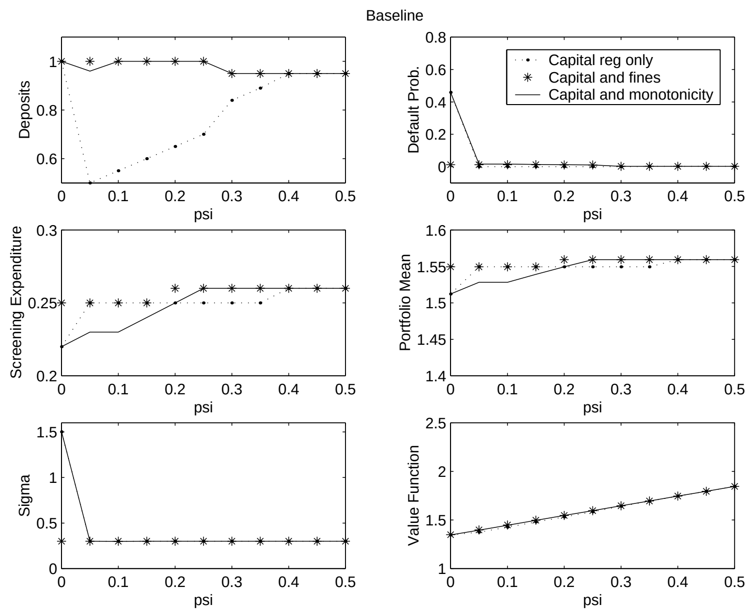

Figure 6: Summary statistics of solutions to three problems: optimal capital regulation (dotted lines), optimal capital regulation with fines ("*"), and optimal capital regulation with fines under the constraint that net income be monotonically increasing in portfolio return (solid lines).

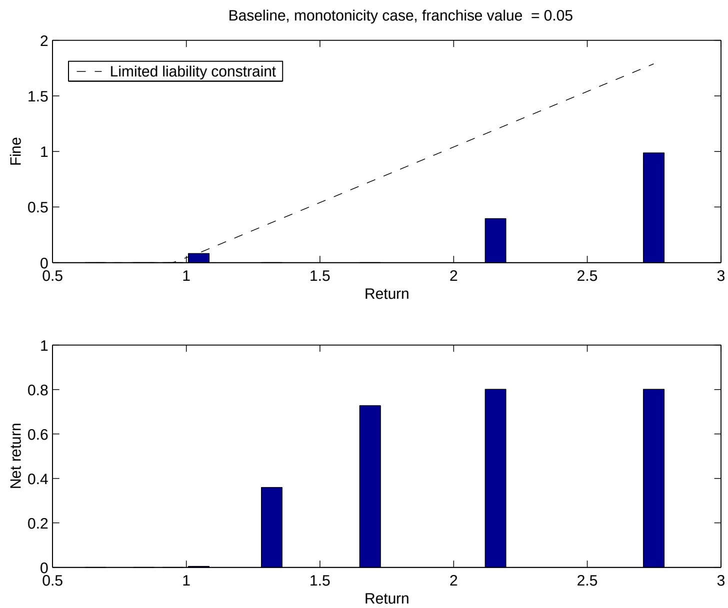

Figure 7: Optimal fine schedule when fines are required to satisfy the monotonicity constraint. Results for baseline paremeterizaiton with $\psi = 0.05$.