

THE WELFARE COST OF BANK CAPITAL REQUIREMENTS

Skander Van Den Heuvel¹

The Wharton School, University of Pennsylvania

May, 2005

Abstract

This paper measures the welfare cost of bank capital requirements and finds that the macroeconomic stakes involved with the design of capital adequacy regulation are large. A general equilibrium model with capital accumulation and a preference for liquidity is developed. Banks provide liquidity services by accepting deposits. A capital requirement plays a role, as it limits the moral hazard on the part of banks that is induced by the presence of a deposit insurance scheme. However, *ceteris paribus*, a higher capital requirement implies that banks can accept fewer deposits and thus provide fewer liquidity services to households. A key result is that equilibrium asset returns reveal the strength of households' preferences for liquidity and it is shown that this allows for the derivation of a simple formula for the welfare cost of capital requirements that is a function only of observable variables. Using U.S. banking data, the cost of increasing the capital requirement by 10 percentage points is equivalent to a permanent loss in consumption of at least 0.1 to 0.2 percent.

¹ Finance Department, Wharton School, 3620 Locust Walk, Philadelphia, PA 19104, USA. Email: vdheuvel@wharton.upenn.edu. The author has benefited from helpful comments from Andy Abel, Joao Gomes, Gary Gorton and seminar participants at Wharton and the Federal Reserve Board. All mistakes are mine.

This paper asks, and provides an answer to, the following question: How large are the welfare costs of bank capital requirements? While there are a number of papers on the theoretical *benefits* of capital adequacy regulation, based on limiting the moral hazard involved with deposit insurance² or externalities associated with bank failures, much less is known about whether there are also costs involved with imposing restrictions on the capital structure of banks. Recent work by Diamond and Rajan (2000) and Gorton and Winton (2000), among others, suggests that capital requirements may have an important cost in that they reduce the ability of banks to create liquidity by accepting deposits. After all, a capital requirement limits the fraction of bank assets that can be financed by issuing deposit-type liabilities. Unfortunately, the models in these papers do not easily lend themselves to a quantification of this cost.

This paper's contribution is to model the benefit of liquidity creation in a flexible way, following Sidrauski (1967), embed the role of liquidity creating banks in an otherwise standard general equilibrium growth model, and use that framework to quantify the welfare cost of increasing the capital requirement. This cost depends crucially on households' preferences for liquidity. A key insight from the model is that equilibrium asset returns reveal the strength of these preferences for liquidity and this allows us to quantify the welfare cost. The model also incorporates a rationale for the existence of capital adequacy regulation, based on a moral hazard problem associated with deposit insurance. The resulting benefits are characterized and, though they harder to quantify, the last section of the paper contains an attempt to measure these as well.

In many countries capital adequacy regulation is currently based on the Basel Accords. In response to perceived shortcomings in the original Accord, practitioners have added more and more detailed refinements, culminating in the soon-to-be implemented Basel 2, while attempting to keep the required ratio of capital to risk weighted assets for a typical bank approximately the same. But is the 8% of the original Basle Accord a good number for the total risk-based capital ratio? This fundamental question remains unaddressed in the literature. At the same time the changes involved in Basel 2 seem likely to greatly increase the cost of compliance and supervision of banks.

² See, for example, Giammarino, Lewis, and Sappington (1993) and Dewatripont and Tirole (1994). See Allen and Gale (2003) for a more skeptical view. Diamond and Dybvig (1983) is often viewed as a theoretical justification for deposit insurance.

If we find that the welfare cost of capital requirements are trivial, this could be an argument for creating a simple, robust system of capital adequacy regulation, with low compliance and supervision costs, but with relatively high capital ratios so as to make bank failure a sufficiently infrequent event. On the other hand, if we find a high welfare cost of capital requirements, this could be an argument for lowering them, by either accepting a higher chance of bank failure, or by designing a more risk-sensitive system with the associated increased supervision and compliance costs, which seems to be the trend in practice.

1. The model

The most important respect in which the model deviates from the standard growth model is that households have a need for liquidity, and that certain agents, called banks, are able to create financial assets, called deposits, which provide liquidity services. Since a central goal of the model is to provide a framework not just for illustrating, but for actually measuring the welfare cost of capital requirements, it is important to model the preferences for liquidity in a way that is not too restrictive. In as much as possible, we would like the data to provide the answer, not the specific modeling choices. To that end, I follow Sidrauski (1967) in adopting the modeling device of putting liquidity services in the utility function.³ This has two disadvantages and one advantage.

One disadvantage is that it does not further our understanding of *why* households like liquid assets, but this is not the topic of this paper, so this concern can be dismissed.⁴ A second disadvantage is that if one needs to specify a particular functional form for the utility function, one is on loose grounds. For example, is the marginal utility of consumption increasing or decreasing in deposits?

Fortunately – and this is the advantage of this approach – there is no need to make unpalatable assumptions of this kind. I will show that it is possible to derive a first-order approximation the welfare cost of raising the capital requirement without making *any* assumptions on the functional form of the utility

³ Lucas (2000) uses the framework of the Sidrauski model to measure the welfare cost of inflation.

⁴ Nonetheless, see Feenstra (1986) for how optimizing models of money demand based on a Baumol Tobin transaction technology can be approximately rewritten as maximization problems with money in the utility function.

function, beyond the standard assumptions that it is increasing and concave. A trade-off involved with modeling liquidity in this flexible way, and embedding it in a general equilibrium analysis, is that the modeling of the banks' assets is not rich enough to incorporate the details of *risk-based* capital requirements.

The environment and the agents' decision problems

Time is discrete and there are infinitely many periods. The economy consists of households, banks, (nonfinancial) firms, and a government or regulatory agent. Households own both the banks and the nonfinancial firms. These firms combine capital and labor to produce the single good which households consume. I now discuss the assumptions for each of these agents, and analyze their decision problems in turn. For the reader's convenience, figure 1 displays a timeline of the model.

Households: There is a continuum of households with mass one. Households are infinitely lived dynasties and have identical preferences. They value consumption and liquidity services. Households can obtain these liquidity services by allocating some of their wealth to bank deposits, an asset created by banks for this purpose. As mentioned, the liquidity services of bank deposits are modeled by assuming that the household's utility function is increasing in the amount of deposits.

Besides holding bank deposits, denoted d_t , households can store their wealth by buying and selling shares, or equity, e_t . They supply a fixed quantity of labor, normalized to one, for a wage, w_t . Taxes are lump-sum and equal to T_t . There is no aggregate uncertainty, so the representative household's problem is one of perfect foresight:

$$\begin{aligned}
 & \max_{\{c_t, d_t, e_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t, d_t) \\
 \text{s.t.} \quad & d_{t+1} + e_{t+1} + c_t = w_t 1 + R_t^D d_t + R_t^E e_t - T_t \\
 & \lim_{T \rightarrow \infty} \left(\prod_{s=0}^T R_s^E \right)^{-1} (d_T + e_T) \geq 0 \\
 & d_0 + e_0 \equiv a_0 \quad \text{given}
 \end{aligned}$$

where c_t is consumption in period t , R_t^D is the return on bank deposits, R_t^E is the return on (bank or firm) equity, and β is the subjective discount factor. The returns R_t^D and R_t^E , and the wage are determined competitively, so the household takes these as given. The same applies for the taxes. There is no distinction between bank and nonbank equity, since, in the absence of risk, they are perfect substitutes for the household and will thus also command the same return. The second constraint is a no-Ponzi game condition, the third an initial condition.

The utility function is assumed to be concave, at least once continuously differentiable on \mathbb{R}_{++}^2 , increasing in both arguments, and strictly increasing in consumption:

$$u_c(c, d) \equiv \partial u(c, d)/\partial c > 0 \quad \text{and} \quad u_d(c, d) \equiv \partial u(c, d)/\partial d \geq 0$$

The first-order conditions to the household's problem are:

$$(c) \quad u_c(c_t, d_t) = \lambda_t$$

$$(d) \quad u_d(c_t, d_t) + \lambda_t R_t^D - \beta^{-1} \lambda_{t-1} = 0$$

$$(e) \quad \lambda_t R_t^E - \beta^{-1} \lambda_{t-1} = 0$$

where λ_t is the Lagrange multiplier associated with the intertemporal budget constraint. Rewriting this,

$$R_t^E = \beta^{-1} (u_c(c_{t-1}, d_{t-1})/u_c(c_t, d_t)) \quad (1)$$

$$u_d(c_t, d_t) = u_c(c_t, d_t) (R_t^E - R_t^D) \quad (2)$$

Equation (1), which determines the return on equity, is the standard Euler equation for the intertemporal consumption-saving choice in a deterministic setting, with one difference: the marginal utility of consumption may depend on the level of deposits. Equation (2) relates the spread between the return on equity and the return on bank deposits to the marginal value of the liquidity services provided by deposits, expressed in units of the consumption good. If $u_d(c, d) > 0$, then the return on equity will be higher than the return on deposits to compensate for the fact that equity does not provide any liquidity services.

Banks: There is a continuum of banks with mass one, which make loans to nonfinancial firms and finance these loans by accepting deposits from households and issuing equity. The ability of banks to create liquidity through deposit contracts is their defining feature. All contracts are resolved in one period. Banks last for one period and every period new banks are set up with free entry into banking. The balance sheet, and the notation, for the representative bank during period t is:

Assets		Liabilities	
L_t	Loans	D_t	Deposits
		E_t	Equity

Banks are subject to regulation, as well as supervision, by the government. One form of regulation is deposit insurance. If a bank fails, the government (through a deposit insurance fund) ensures that no depositor suffers a loss as a consequence of this failure. That is, all deposits are fully insured. Equity holders, as residual claimants, are left with nothing in the event of failure. The rationale for the deposit insurance is left unmodeled. However, it has been argued that deposit insurance improves the ability of banks to create liquidity.⁵

Secondly, banks face a capital requirement, which requires them to have a minimum amount of equity as a fraction of (risk-weighted) assets. Since loans are the only type of asset in this model, the capital requirement simply states that equity needs be at least a fraction γ of loans for a bank to be able to operate:

$$E_t \geq \gamma L_t$$

For the moment, the capital requirement is merely assumed. It will later be shown how it can be socially desirable to have such a requirement, as it mitigates the moral hazard problem that arises due to the presence of deposit insurance.

The bank can make loans to nonfinancial firms, described below. These loans are riskless⁶ and yield a rate of return R_t^L . R_t^L is determined competitively in equilibrium, so each bank takes it as given. Thus, a bank that lends out L_t units of the good at the beginning of the period will receive nonrandom total return of

⁵ Diamond and Dybvig (1983) provide a model of panic based bank runs, which can be seen as a rationale for deposit insurance.

⁶ The assertion that the bank can make riskless loans is a consequence of the technology of the nonfinancial firms, as detailed below.

$R_t^L L_t$ units at the end of the period. It is assumed that there are no transaction costs involved in making loans and accepting deposits. Section 4 will analyze an extension of the model with costly financial intermediation.

The presence of deposit insurance creates a moral hazard problem: the bank has an incentive to engage in excessive risk-taking. Since this is the justification for the capital requirement, I introduce a way for the bank to engage in excessive risk-taking by assuming that the bank has the option of artificially raising the riskiness of its assets. Specifically, by lending to a different set of nonfinancial firms with a risky technology,⁷ the bank can create a loan portfolio with riskiness σ that pays off $R_t^L + \sigma_t \varepsilon_t$, where ε_t is a bank-specific shock with the following distribution:

$$\varepsilon_t = \begin{cases} 1 & \text{with probability 0.5} \\ -(1 + 2\xi) & \text{with probability 0.5} \end{cases} \quad (3)$$

where ξ is a positive constant. It is assumed that the bank can choose the riskiness of its loans, though, as will be explained in the discussion of the government, bank supervision imposes an upper bound on the choice: $\sigma_t \in [0, \bar{\sigma}]$. The variance of the return is $\sigma_t^2 (1 + \xi)^2$, which is increasing in σ_t , and its expected value is $R_t^L - \sigma_t \xi$, which is decreasing in σ_t . It is in this sense that risk-taking is excessive: absent a moral hazard problem due to deposit insurance, the bank would always choose $\sigma_t = 0$. The very special assumption (3) is made purely for expositional reasons. All the results in this paper hold for an arbitrary distribution of ε with bounded support and nonpositive mean (see Appendix D for a proof).

I am now in a position to state the bank's problem. The objective of the bank is to maximize shareholder value by deciding how many loans to make, how much risk to take on, and how to finance its assets with equity and deposits. Although the decision on how much equity to issue will be endogenized, it is convenient to first analyze the sub-problem of maximizing shareholder value right after the equity has been issued and the bank has raised E_t in equity at the beginning of period t . At that point the value of the bank to its shareholders is:⁸

⁷ Again, the technology will be described below and will be consistent with the rates of return assumptions made here.

⁸ In what follows, time subscripts will be used only where necessary to avoid confusion.

$$\begin{aligned}
V^B(E) = \max_{L,D,\sigma} & \mathbb{E}_\varepsilon \left[\left((R^L + \sigma\varepsilon)L - R^D D \right)^+ \right] / R^E \\
\text{s.t.} & \quad L = E + D \\
& \quad E \geq \gamma L \\
& \quad \sigma \in [0, \bar{\sigma}]
\end{aligned} \tag{4}$$

The notation $(x)^+$ stands for $\max(x, 0)$ and \mathbb{E} is the expectations operator. The first constraint is the balance sheet identity, the second is the capital requirement, and the third bounds σ .

The term $(R^L + \sigma\varepsilon)L - R^D D$ is the bank's net cash flow at the end of the period. It consists of interest income from loans, minus any possible charge-offs on the loans, and minus the interest owed to depositors. If the net cash-flow is positive, shareholders are paid the full amount in dividends. If the net cash flow is negative, the bank fails and the deposit insurance fund must cover the difference in order to indemnify depositors, as limited liability of shareholders rules out negative dividends. Shareholders receive zero in this event, so dividends equal $\left((R^L + \sigma\varepsilon)L - R^D D \right)^+$.

At the beginning of period t shareholders discount the value of dividends, which are paid at the end of that period, by their opportunity cost of holding this particular bank's equity. This opportunity cost is R^E , the market rate of return on equity. Because dividends are either not subject to risk, or, if $\sigma > 0$, their risk is perfectly diversifiable, shareholders do not price the bank's risk.⁹

First, I characterize the choice of σ conditional on L and D . Note that

$$\begin{aligned}
\mathbb{E}_\varepsilon \left[\left((R^L + \sigma\varepsilon)L - R^D D \right)^+ \right] \\
= \begin{cases} (R^L - \sigma\xi)L - R^D D & \text{if } (R^L - \sigma(1+2\xi))L - R^D D \geq 0 \\ 0.5((R^L + \sigma)L - R^D D) & \text{otherwise} \end{cases}
\end{aligned}$$

Expected dividends are thus strictly decreasing in σ for low values of σ and strictly increasing in σ for sufficiently high values of σ .¹⁰ The reason is that for high values of σ , if the bank suffers a negative shock, there is not enough equity to absorb the loss and the excess loss is covered by the deposit insurance fund. Increasing risk further at this point increases the payoff to shareholders in the

⁹ Hence, the treatment of R^E as nonstochastic in the household problem is still correct, since, even if banks are risky, households would not leave any such risk undiversified.

¹⁰ Note that there is *no* discontinuity at $(R^L - \sigma(1+2\xi))L - R^D D = 0$.

good state ($\varepsilon = 1$) without lowering it in the bad state. In other words, the value of the put option associated with the deposit insurance fund increases with σ . In contrast, when σ is low, the value of this put option is zero and shareholders fully take into account the reduction in net present value that occurs when risk is increased.

Because expected dividends are a convex function of σ , there are only two values to consider for the optimal choice of riskiness: $\sigma = 0$ or $\sigma = \bar{\sigma}$. It is easy to show that

$$\begin{aligned}\sigma = 0 &\text{ iff } \bar{\sigma} \leq R^L - R^D(D/L) \\ \sigma = \bar{\sigma} &\text{ otherwise}^{11}\end{aligned}\tag{5}$$

Because $E = L - D \geq \gamma L$, the following is a sufficient condition for $\sigma = 0$:

$$\bar{\sigma} \leq R^L - R^D(1 - \gamma)\tag{6}$$

This is also a necessary condition when the capital requirement is binding. From now on, unless explicitly stated otherwise, it is assumed that (6) holds.

The bank's sub-problem in (4) now simplifies to:

$$\begin{aligned}V^B(E) &= \max_{L,D} (R^L L - R^D D) / R^E \\ \text{s.t.} \quad L - D &= E \\ E - \gamma L &\geq 0\end{aligned}$$

The first-order conditions are easily simplified to

$$R^L - R^D = \gamma R^E \chi$$

where χ is the Kuhn-Tucker multiplier associated with the capital requirement: $\chi \geq 0$ and $\chi(E - \gamma L) = 0$. The existence of a (finite) solution requires $R^L \geq R^D$. Under that condition, the solution is

$$V^B(E) = (R^L + (\gamma^{-1} - 1)(R^L - R^D))E / R^E\tag{7}$$

¹¹ When $\bar{\sigma} = R^L - R^D(D/L)$, the bank is indifferent between the two choices. For convenience, it is assumed that the bank chooses $\sigma = 0$ in that case.

The capital requirement binds if and only if $R^L > R^D$. The interpretation is straightforward: an extra unit of equity can be lent out at the rate R^L . In addition, the extra unit of capital allows the bank to make $(\gamma^{-1} - 1)$ *additional* loans and finance those with deposits, without violating the capital requirement, which requires $L \leq \gamma^{-1}E$. If $R^L > R^D$, the second option has value, and the capital requirement will be binding, otherwise not.

I can now turn to the bank's decision on how much equity to raise. The pre-issue value of the bank is $V^B(E) - E$. The bank maximizes this value when choosing E :

$$\begin{aligned}\pi^B &= \max_E (V^B(E) - E) \\ &= \max_E (R^L + (\gamma^{-1} - 1)(R^L - R^D))E / R^E - E\end{aligned}$$

The first-order condition is:¹²

$$R^E = R^L + (\gamma^{-1} - 1)(R^L - R^D)$$

Since we have already established that $R^L \geq R^D$ is necessary for a solution to exist and that the capital requirement binds if and only if $R^L > R^D$, we can distinguish two cases:

1. If $R^L = R^D$, the capital requirement is slack and

$$R^E = R^L = R^D \tag{8}$$

2. If $R^L > R^D$, the capital requirement is binding, so $E = \gamma L$, and

$$R^L = (1 - \gamma)R^D + \gamma R^E \tag{9}$$

In either case, $V^B(E) - E = 0$.

In the case of a binding capital requirement, one unit of lending is financed by γ in equity and $(1 - \gamma)$ in deposits, so for zero profits the rate of return to lending better

¹² As is common in problems with constant returns to scale, the first order condition, rather than fully determining the agent's choice, has the interpretation of a necessary condition for the existence of a finite solution. If $R^E < (>)R^L + (\gamma^{-1} - 1)(R^L - R^D)$, then E tends to plus (minus) infinity. If the first-order condition holds, E is indeterminate, and thus so is the scale of the bank.

equal the similarly weighted average of the required rates of return of equity and deposits, which is what is stated in (9).

Note that the sufficient condition for $\sigma = 0$ to be optimal, given in (6), is seen to be equivalent to

$$\bar{\sigma} \leq \gamma R^E \quad (10)$$

Again, this condition is also necessary if the capital requirement is binding.

Firms: Nonfinancial firms cannot create liquidity through deposits. They can, however, buy goods to use them as capital, which can be combined with labor input, to produce output of the good. Capital is purchased at the beginning of the period. To finance their capital stock, firms can issue equity to households, borrow from banks, or some combination of both. The firm's balance sheet, and notation, for period t is:

Assets		Liabilities	
K_t	Physical Capital	L_t	Loans
		E_t^F	Firm Equity

There is a continuum of firms with mass normalized to one. Firms can employ a riskless or a risky production technology. The riskless technology is standard.¹³ Output in period t is $F(K_t, H_t)$, where H_t is hours of labor input and $F(\cdot)$ is a well-behaved production function exhibiting constant returns to scale. A fraction δ of the capital stock depreciates during the period. There are no adjustment costs. The absence of adjustment costs implies that firm's problem is static.

As in the analysis of the bank's problem, it is convenient to start with the firm's decision problem right after it has raised E_t^F in equity. At that point the value of the firm to its shareholders is

$$V^F(E^F) = \max_{K, H} \left(F(K, H) + (1 - \delta)K - wH - R^L(K - E^F) \right) / R^E$$

¹³ It would be straightforward to let both types of firms have risky production, and therefore, risky loans to individual firms even while keeping $\sigma = 0$ as feasible for banks, as long as the production shocks are sufficiently imperfectly correlated across firms, so that the risk is perfectly diversifiable by lending to many firms. Excessive risk taking would then correspond to not diversifying this risk.

Here I have substituted out loans using the balance sheet identity.¹⁴ The first-order conditions for the choices of capital and labor inputs are standard:

$$(H) \quad F_H(K, H) = w \quad (11)$$

$$(K) \quad F_K(K, H) + (1 - \delta) = R^L \quad (12)$$

These optimality conditions, together with the constant returns to scale assumption, imply that the solution for the firm's shareholder value is: $V^F(E^F) = (R^L E^F) / R^E$. The pre-issue value of the firm is $V^F(E^F) - E^F$. The firm maximizes this value when choosing E^F . It is assumed that equity cannot be negative.

$$\pi^F = \max_{E^F \geq 0} V^F(E^F) - E^F$$

The first order condition is

$$(E^F) \quad R^E = R^L + \mu, \quad \mu \geq 0, \quad \mu E^F = 0 \quad (13)$$

where μ is the Kuhn-Tucker multiplier associated with the constraint that firm equity cannot be less than zero.¹⁵ A finite solution thus requires $R^L \leq R^E$.

If $R^L < R^E$, then $E^F = 0$, so $K = L$. In other words, if bank loans are cheaper than equity finance, the firm chooses to use only bank loans to finance its capital. If $R^E = R^L$, the firm's financial structure is not determined by individual optimality. In either case economic profits, π^F , equal zero.

Which case applies turns out to depend on whether or not the demand for liquidity is satiated. To see this, recall that the analysis of the household's problem established that the required rate of return for risk-free equity is higher than the required rate of return for bank deposits: $R^D \leq R^E$ (see equations (1) and (2)). Moreover, the inequality is strict if and only if the households' demand for liquidity is nonsatiated. From the bank's problem we know that $R^L = (1 - \gamma)R^D + \gamma R^E$.¹⁶ It follows that, in fact, $R^L \leq R^E$ and that this inequality is strict if and only if the households' demand for liquidity is nonsatiated.

¹⁴ Note that the absence of arbitrage opportunities implies that nonfinancial firms have to offer shareholders the same return on equity as banks, since there is no aggregate risk.

¹⁵ Technically, μ is the multiplier to that constraint after rescaling the problem by R^E , a constant to the firm.

¹⁶ Note that this is true whether the capital requirement slack or binding.

Hence, as long as liquidity preference is not satiated, nonfinancial firms finance all their capital stock with the cheaper bank loans, rather than equity. If instead the demand for liquidity is satiated ($u_d(c, d) = 0$), firms are indifferent between the two forms of finance. For this less interesting case, I assume without loss of generality that firms still opt for bank loans only.¹⁷ Hence, in both cases,

$$E^F = 0 \quad (14)$$

In addition to this riskless technology, firms can also employ a risky technology. Risky firms provide a vehicle for banks to make the kind of risky loans described in the subsection on banks. Although these firms thus provide a rationale for the existence of capital regulation, as mentioned, I will usually focus on the case that the capital requirement is sufficiently high, according to condition (10), to prevent banks from engaging in excessive risk taking. These risky firms will then not exist in equilibrium. For this reason analysis of these firms is left for appendix A. Here I simply state the assumptions regarding the risky technology.

Risky firms differ in that their output is subject to an additive shock that is proportional to their capital stock: output is $F(K, H) + \sigma_{RF} \varepsilon K$, where ε is the same negative mean shock as defined in (3) and σ_{RF} is a parameter ($\sigma_{RF} \geq \bar{\sigma}$). The choice of technology (i.e. the type of firm) is observable to all parties to a financial contract with the firm, as is the value of ε when realized. The appendix shows how the optimal loan contract with such a firm gives a bank the ability to create a loan portfolio with riskiness σ by directing a fraction σ/σ_{RF} of lending to risky firms with perfectly correlated shocks (e.g. by lending to one risky firm).

Government: The government manages the deposit insurance fund, sets a capital requirement $\gamma \in [0, 1)$ and conducts bank supervision. Bank supervision is necessary to enforce the capital requirement. A second role for bank supervision is to monitor excessive risk taking by banks, σ . Supervisors can to some degree detect such behavior and stop any bank that is ‘caught’ attempting to take on excessive risk in order to protect the deposit insurance fund. It seems reasonable to assume that a small amount of risk taking is harder to detect than a large

¹⁷ This assumption is made for convenience and is without loss of generality in the sense that it has no effect on the model’s predictions for any equilibrium rates of return, the paths of consumption and capital accumulation, or the measurement of the welfare cost of the bank capital requirement.

amount. The largest level of risk-taking that is still just undetectable is $\bar{\sigma}$. $\bar{\sigma}$ is assumed to be a decreasing function of the resources devoted to bank supervision:

$$\bar{\sigma} = S(T) \text{ with } S'(\bullet) \leq 0 \text{ and } 0 < S \leq \sigma_{RF}$$

where T , a choice variable for the government, is the part of tax revenue spent on bank supervision.¹⁸ The interpretation is that, as more resources are devoted to bank supervision, banks are less able to engage in excessive risk taking without being detected.

The assumption that excessive risk taking is *imperfectly* observable is important. If regulators could perfectly observe each bank's riskiness, they could simply adjust each bank's deposit insurance premium so as to make the bank pay for the *ex ante* expected loss to the deposit insurance fund, thus eliminating any moral hazard. Or they could set each bank's capital requirement as an increasing function of its riskiness, so as to ensure that the bank always internalizes all the risks. But such perfect observability is simply not realistic, so a moral hazard problem does exist. Not allowing σ to exceed $\bar{\sigma}$ can be interpreted as a risk-based capital requirement or a risk-based deposit insurance premium, but one based on *observable* risk: regulators may deter detectable excessive risk taking (as they should) by imposing a sufficiently high capital requirement or a sufficiently high deposit insurance premium when such excessive risk taking is detected.¹⁹

The government has a balanced budget. Lump-sum taxes are set at

$$T_t = T + 0.5 \left(R_t^D D_t - (R_t^L - \sigma_t(1 + 2\xi)) L_t \right)^+. \quad (15)$$

The second term represents the cost reimbursing depositors of failed banks under the deposit insurance scheme. If (10) holds, we know that $\sigma_t = 0$ and in that case taxes are simply: $T_t = T$.

¹⁸ As in the standard growth model with government spending and lump sum taxes, if T is set 'too high', no equilibrium with positive consumption will exist. I assume that T is sufficiently low so that a steady state equilibrium with positive consumption exists. Appendix B makes precise what 'sufficiently low' means for a particular functional form of the utility function.

¹⁹ As long as it is sufficiently high, the precise value of the capital requirement or premium when $\sigma > \bar{\sigma}$ is not important, since it will never be implemented in equilibrium!

General Equilibrium

Given a government policy γ and T , an equilibrium is defined as a path of consumption, capital, deposits, equity holdings, bank loans and financial returns, for $t = 0, 1, 2, \dots$ such that:

1. Households, banks and nonfinancial firms all solve their maximization problems, described above, with taxes set according to (15);
2. All markets clear, i.e.

$$\begin{aligned} d_t &= D_t \\ e_t &= E_t + E_t^F \\ L_t &= K_t - E_t^F \\ H_t &= 1 \end{aligned}$$

and

$$F(K_t, 1) - \xi \sigma_t L_t + (1 - \delta) K_t = c_t + K_{t+1} + T;$$

I focus on the case that (10) holds: $S(T) \leq \gamma R_t^E$. Government policy can accomplish this by setting γ and/or T sufficiently high. In that case, $\sigma_t = 0$ and $T_t = T$. Combining the market clearing conditions and equations (1), (2), (8), (9), (11), (12), and (14) it is possible to characterize the equilibrium in terms of a system in (K_t, c_t) with R_t^E and d_t as auxiliary variables:

$$K_t = F(K_{t-1}, 1) + (1 - \delta) K_{t-1} - c_{t-1} - T \quad (16)$$

$$\beta^{-1}(u_c(c_{t-1}, d_{t-1}) / u_c(c_t, d_t)) = R_t^E \quad (17)$$

$$F_K(K_t, 1) + 1 - \delta = R_t^E - (1 - \gamma) \frac{u_d(c_t, d_t)}{u_c(c_t, d_t)} \quad (18)$$

where d_t is determined according to one of the following two cases:

1. If $u_d(c_t, (1 - \gamma)K_t) = 0$, the capital requirement is not binding and

$$u_d(c_t, d_t) = 0 \text{ with } d_t \leq (1 - \gamma)K_t \quad (19)$$

2. If $u_d(c_t, (1 - \gamma)K_t) > 0$, the capital requirement is binding and

$$d_t = (1 - \gamma)K_t \quad (20)$$

Remark: $L_t = K_t$, $e_t = E_t = K_t - d_t$, and remaining variables are determined through (2), (11) with $H_t = 1$, and (8). In case 1, which requires that demand for liquidity

be satiated at $d = (1 - \gamma)K_t$, d_t and e_t are not uniquely determined. Note that if $u_d(c_t, d_t) = 0$, $u_c(c_t, d_t)$ does not depend on d_t .

In the first case, the capital requirement is nonbinding, because the level of deposits is such that the marginal value of liquidity provision is zero, which implies that their rate of return is no different from equity: $R_t^D = R_t^E$. Except for the lump sum taxes, there is no material difference here with a standard growth model. Banks channel funds from households to firms, but they do not really matter at the margin, since their special ability to create liquidity has no marginal value ($R_t^L = R_t^E$).

In the second case, the pecuniary return on deposits is lower than the return on equity by a spread equal to $u_d(c_t, d_t)/u_c(c_t, d_t)$, which is the marginal value of deposits' liquidity services expressed in units of consumption. The capital requirement is now binding, since banks want to fund their assets as much as possible with the cheaper deposits. Because banking is characterized by perfect competition, banks fully pass on the lower cost of funding their loans to their borrowers, firms. However, since banks still have to finance a fraction γ of their lending with equity, the loan rate declines only by $(1 - \gamma)(u_d(c_t, d_t)/u_c(c_t, d_t))$.

A steady state result: No superneutrality

Because banks pass on the low pecuniary return on deposits to their borrowers, the steady state level of the capital stock is not generally invariant to capital adequacy regulation. This contrasts with the superneutrality result of the Sidrauski (1967) model.²⁰ In the present model, raising γ can increase or lower the steady state capital stock, depending on the interest elasticity of liquidity demand. For example, suppose that

$$u(c, d) = \tilde{u}(\phi(c, d)) \text{ and } \phi(c, d) = \left\{c^{(\eta-1)/\eta} + ad^{(\eta-1)/\eta}\right\}^{\eta/(\eta-1)}$$

with $a, \eta > 0$, $\tilde{u}' > 0$ and $\tilde{u}'' < 0$. Then

²⁰ In the Sidrauski model, the rate of inflation (which is what determines the rate of return on monetary assets) has no impact on the steady state capital stock. The reason is that in that model money is created by a monopolistic entity, the government, which does not in any way use the revenues from liquidity creation (seigniorage) to lower the marginal cost of funding investment.

$$\frac{u_d(c_t, d_t)}{u_c(c_t, d_t)} = a \left(\frac{d_t}{c_t} \right)^{-1/\eta},$$

so that demand for deposits is given by

$$d_t = a^\eta c_t (R_t^E - R_t^D)^{-\eta}.$$

Hence, η has the interpretation of the interest elasticity of the demand for deposits. It is straightforward to show that for this specification, *the steady state level of the capital stock, K^* , is increasing (decreasing) in γ if $0 < \eta < 1$ ($\eta > 1$)*. A proof can be found in Appendix B.²¹

The intuition for this result is as follows. Firms set the marginal product of capital equal to the rate on bank loans, which in turn equals

$$R^L = R^E - (1 - \gamma)(R^E - R^D).$$

(cf. (9)). In steady state, $R^E = \beta^{-1}$. An increase in γ has two effects on R^L : one is to force banks to rely more on equity finance, which is more expensive than deposits ($R^E - R^D > 0$). This effect, which is explicit in the above equation, increases R^L . The second effect is a general equilibrium feedback. The fact that bank must rely less on deposits makes them more scarce to households, which increases the spread $R^E - R^D$. This second effect lowers the competitive rate on bank loans. If the interest elasticity of the demand for deposits is low ($0 < \eta < 1$), a large increase in the spread will be necessary to convince households to make do with fewer deposits, and the second effect will dominate. In that case, R^L falls and the steady state level of capital thus rises, otherwise not.

A related point is that the steady state level of the capital stock depends on the strength of the preference for liquidity. Without any preference for liquidity, equity, deposits and bank loans all command the same return, equal to β^{-1} in steady state. In the above example, this occurs when $a = 0$. On the other hand, if liquidity has value ($a > 0$ in the example), the steady state return on equity will still be β^{-1} , but the return on deposits, and therefore the return on bank loans, will be strictly lower than that. Since firms equate the marginal product of capital

²¹ The proof in Appendix B makes explicit the assumption that taxes T are sufficiently low to guarantee the existence of a steady state equilibrium with positive consumption. See footnote 18.

to the rate on bank loans, it follows that the steady state capital stock is higher when households exhibit a preference for liquidity.

2. The welfare cost of the capital requirement: a first-order approximation

The strategy for quantifying the welfare cost of the capital requirement is as follows. First, I present the constrained social planner's problem to the model. The qualification 'constrained' means that the social planner's problem shall respect the capital requirement and devote the same level of resources to supervision. This will ensure that the allocation that solves the social planner's problem is incentive compatible for the banks. I then show that the allocation associated with the social planner's problem is identical to the allocation of the decentralized equilibrium described above. Finally, I use this equivalence to derive analytically a first-order approximation of the welfare cost of increasing the capital requirement γ .

The social planner's problem

Define the following constrained social planner's problem:

$$\begin{aligned}
 V_0(\theta) = & \max_{\{c_t, d_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t, d_t) \\
 \text{s.t.} \quad & F(K_t, 1) + (1 - \delta)K_t = c_t + K_{t+1} + T \\
 & (1 - \gamma)K_t - d_t \geq 0
 \end{aligned} \tag{21}$$

where $\theta = (\gamma, T, \delta, \beta, K_0)$. The first constraint is the social resource constraint for $\sigma = 0$,²² the second constraint rewrites the capital requirement. The first-order conditions to this problem are:

$$(c) \quad u_c(c_t, d_t) = \lambda_t^{sp}$$

$$(d) \quad u_d(c_t, d_t) = \chi_t^{sp}$$

²² The absence of excessive risk taking is simply part of the definition of the constrained social planner's problem.

$$(K) \quad \lambda_t^{sp}[F_K(K_t, 1) + 1 - \delta] - \beta^{-1} \lambda_{t-1}^{sp} + \chi_t^{sp}(1 - \gamma) = 0$$

where λ_t^{sp} and χ_t^{sp} are the Lagrange multipliers on the social resource constraint and the capital requirement, respectively. Combining these first-order conditions yields:

$$F_K(K_t, 1) + 1 - \delta = \left(\frac{\beta^{-1} u_c(c_{t-1}, d_{t-1})}{u_c(c_t, d_t)} \right) - (1 - \gamma) \left(\frac{u_d(c_t, d_t)}{u_c(c_t, d_t)} \right) \quad (22)$$

In addition, since $u_{dd}(c, d) \leq 0$ by concavity of u , the first order condition with respect to deposits, combined with the complementary slackness conditions $d_t \leq (1 - \gamma)K_t$, $\chi_t^{sp} \geq 0$ and $\chi_t^{sp}((1 - \gamma)K_t - d_t) = 0$, implies :

$$\text{if } u_d(c_t, (1 - \gamma)K_t) = 0 \text{ then } d_t \leq (1 - \gamma)K_t, \text{ with } u_d(c_t, d_t) = 0; \quad (23)$$

$$\text{if } u_d(c_t, (1 - \gamma)K_t) > 0 \text{ then } d_t = (1 - \gamma)K_t \quad (24)$$

Combining equations (22), (23) and (24) with the social resource constraint (the first constraint to the problem in (21)), it is apparent that the allocations of K_t , c_t and d_t are identical to those of the decentralized equilibrium summarized above in equations (16) through (20). Equation (23) corresponds to an equilibrium with a nonbinding capital requirement ('case 1', equation (19)), while equation (24) corresponds to the case of a binding capital requirement ('case 2', equation (20)).

Hence, *the constrained social planner's problem replicates the decentralized equilibrium* when $\sigma = 0$ in the decentralized equilibrium. As a result, if $\sigma = 0$, *welfare in the decentralized equilibrium is equal to $V_0(\theta)$* , the value of the objective function to the constrained social planner's problem.

A formula for the marginal welfare cost

The equivalence of the constrained social planner's problem and the decentralized equilibrium can be used to measure the marginal effect on welfare of a change in the capital requirement in the following way. Call the current period period 0. Assume government policy is such that (10) holds: $S(T) \leq \gamma R_t^E$ for all $t \geq 0$, so that $\sigma_t = 0$ in the decentralized equilibrium. That is, the capital requirement is sufficiently high to be successful in preventing excessive risk

taking. Starting from this situation, I compute the *marginal effect on welfare of raising γ , without altering T* , using the envelope theorem, as follows:

$$\frac{\partial V_0(\theta)}{\partial \gamma} = -\sum_{t=0}^{\infty} \beta^t \chi_t^{sp} K_t = -\sum_{t=0}^{\infty} \beta^t u_d(c_t, d_t) K_t$$

The last equality follows from the first order condition (d) to the social planner's problem. Since the allocations of K_t , c_t and d_t are identical to those of the decentralized equilibrium, I can use the decentralized equilibrium values for the variables on the right hand side of this equation. Moreover, in the decentralized equilibrium, we have, using (2),

$$\begin{aligned} u_d(c_t, d_t) K_t &= u_c(c_t, d_t)(R_t^E - R_t^D) K_t \\ &= u_c(c_t, d_t)(R_t^E - R_t^D) d_t / (1 - \gamma) \end{aligned}$$

where the last equality follows from the fact that $d_t = (1 - \gamma)K_t$ if the capital requirement binds, and that $(R_t^E - R_t^D) = 0$ if it does not bind.

I compare this to the welfare effect of a permanent change in consumption by a factor $(1 + \nu)$. Starting from the initial equilibrium, the effect on welfare of changing consumption from c_t to $(1 + \nu)c_t$, for all t , equals to a first-order approximation,

$$\left(\sum_{t=0}^{\infty} \beta^t u_c(c_t, d_t) c_t \right) \nu$$

Next, let's assume that the economy is in steady state in period 0. Then the first order approximation of the welfare effect of an increase in γ by $\Delta\gamma$ simplifies as follows:

$$\frac{\partial V_0(\theta)}{\partial \gamma} \Delta\gamma = -\frac{u_d(c_0, d_0) K_0}{1 - \beta} \Delta\gamma = -\frac{u_c(c_0, d_0)(R_0^E - R_0^D) d_0}{(1 - \beta)(1 - \gamma)} \Delta\gamma$$

while

$$\left(\sum_{t=0}^{\infty} \beta^t u_c(c_t, d_t) c_t \right) \nu = \left(\frac{u_c(c_0, d_0) c_0}{1 - \beta} \right) \nu \quad (25)$$

if period 0 is a steady state. Equating the right hand sides of these last two equations, we have the following result.²³

Proposition 1 *Assume the economy is in steady state in period 0 and (10) holds.*

Consider permanently increasing γ by $\Delta\gamma$. A first-order approximation to the resulting welfare loss, expressed as the welfare-equivalent permanent relative loss in consumption, is

$$\nu(\Delta\gamma) = \frac{d_0}{c_0} (R_0^E - R_0^D) (1 - \gamma)^{-1} \Delta\gamma \quad (26)$$

The above formula is empirically implementable. Remarkably, it does not rely on any assumptions about the functional form of preferences, beyond the standard assumptions of monotonicity, differentiability and concavity. Instead, the formula relies on asset returns to reveal the representative household's preference for liquidity relative to consumption. An unnecessary increase in the capital requirement reduces welfare by reducing the ability of banks to issue deposit-type liabilities for any given level of bank assets.

The first factor in the formula for the welfare loss concerns the importance of deposits in the economy. The second is the spread between the return on bank equity and the pecuniary return to deposits. This spread equals the amount of consumption households are willing to forgo in order to enjoy the liquidity services of one additional unit of deposits. Finally, $(1 - \gamma)^{-1} \Delta\gamma$ is the relative change in deposits as a result of changing the capital requirement for a given level of bank assets.

Note that, while the proposition assumes that the economy is initially in steady state, the welfare loss takes into account, to a first-order approximation, all the gains or losses associated with the transition to a new steady state upon changing the capital requirement. That is, simply comparing the welfare levels of different steady states associated with different values of γ yields a different (and wrong) answer because it does not take into account the welfare effects of the

²³ Equate the right hand sides of these two equations to solve for ν . For $\Delta\gamma > 0$, the result is a negative value for ν which, to a first-order approximation, consumers would be equally unhappy to experience as a rise in the capital requirement by $\Delta\gamma$. $\nu(\Delta\gamma)$ in the proposition is the absolute value of this negative value, which is thus the first-order approximate welfare-equivalent relative loss in consumption associated with $\Delta\gamma$.

transition between the steady states. (Recall that there is no superneutrality.) Interestingly, the formula can also be ‘derived’ by incorrectly (!) assuming that the equilibrium levels of the capital stock and bank assets are invariant to changes in γ .²⁴ The fact that this is true is a manifestation of the envelope theorem: these quantities are constrained optimal in the sense of the social planner’s problem, so their response to a change in γ has only a second-order effect on welfare. Of course, one would not have known this before going through the entire exercise.

It may still seem surprising that no assumptions were needed on functional form the utility function. After all, to use the Sidrauski model to measure the welfare cost of inflation, as in Lucas (2000), one does need to specify the functional form of the utility function. The difference is that, while money in the Sidrauski model is created by a nonoptimizing monopolist (the government), in this model the supply of liquidity is created by competitive banks. This additional structure in the model means we have some extra information on the welfare effect of the change in the quantity of deposits. In contrast, in the Sidrauski model, to know the change in real balances in response to a change in inflation, one needs to know the interest elasticity of money demand, which amounts to requiring more knowledge of the utility function.

The optimal capital requirement

The rationale for capital adequacy regulation in the model is its role, joint with bank supervision, in preventing excessive risk taking. If bank supervision is imperfect ($S(T) > 0$ for all T) and preventing excessive risk taking is socially optimal, then the optimal capital requirement will be strictly positive. In the present model preventing excessive risk taking can be shown to be socially optimal if its cost ξ is sufficiently large.²⁵ Under that hypothesis, in steady state, the capital requirement that maximizes welfare is defined by:

²⁴ This is not inconsistent with the previous statements: the welfare effect of the transition to a new steady state cancels with the difference in welfare between the two steady states that is due to the different level of K^* rather than γ alone.

²⁵ If ξ is small relative to the strength of liquidity preference, the social optimum is to set $\gamma = 0$ and accept the result that half the banks will fail. In that case, the welfare cost of the capital requirement is a measure of the government’s ‘stupidity’, rather than one side of an equation that determines the optimal capital requirement.

$$\begin{aligned} & \max_{T, \gamma} V_0(\theta) \\ \text{s.t. } & \gamma\beta^{-1} \geq S(T) \end{aligned}$$

The constraint is the incentive compatibility condition (10), with $R^E = \beta^{-1}$ in steady state. The first order conditions to this problem imply

$$\frac{\partial V_0(\theta)}{\partial \gamma} + \left(\frac{\partial V_0(\theta)}{\partial T} \right) \frac{dT}{d\gamma} \Big|_{S(T)=\gamma\beta^{-1}} = 0$$

Evaluating this in steady state yields

$$\nu(1)c_0 = -\frac{dT}{d\gamma} \Big|_{S(T)=\gamma\beta^{-1}} = \frac{-1}{\beta S'(T)} \quad (27)$$

That is, the marginal welfare cost of the capital requirement (in units of the good per period) should equal its marginal benefit in reducing bank supervision and compliance costs, given the incentive compatibility constraint. Making the reasonable assumption that there are diminishing returns to bank supervision, so that $S'' > 0$, a larger welfare cost demands higher supervision expenditures (a larger T) and thus a lower capital requirement. Unfortunately, measuring the marginal benefit of the capital requirement based on this expression seems less than straightforward. This is not true for the marginal welfare cost, and I now turn to quantifying this cost.

3. Measurement of the welfare cost

The main result so far is an expression for the welfare cost of a bank capital requirement. The expression lends itself to a calculation of this cost based on data. The data used here are annual aggregate balance sheet and income statement data for all FDIC-Insured Commercial Banks in the United States (50 states and DC), obtained from the FDIC's Historical Statistics on Banking (HSOB). These data are based on regulatory filings.

In mapping the theory to the data, some choices need to be made. For deposits, $D (= d)$, the HSOB's Total Deposits is used. The net return on deposits

$(R^D - 1)$ is calculated as Interest on Total Deposits divided by Total Deposits.²⁶ For consumption, c , I use personal consumption expenditures from the NIPA. As a measure of the capital requirement γ the empirical counterpart of E/L is used.²⁷ This is computed as Total Equity Capital plus Subordinated Notes divided by Total Assets. Subordinated Notes are included because subordinated debt counts, within certain limits, towards regulatory tier 2 capital. Total Equity Capital plus Subordinated Notes does not exactly correspond to total capital in the sense of the Basle Accord, on which current capital adequacy regulation in the US (and many other countries) is based. However, data on total capital in the sense of the Basle Accord is only available starting in 1996 and it seems more important to be able to use a longer time span, especially since the formula for the marginal welfare cost in (26) is not very sensitive to the measurement of γ .

An alternative would have been to use the actual regulatory numbers for the capital requirement (either 0.08 for total capital based on the Basle Accord or 0.10 based on the FDICIA, the CAMELS ratings and the Gramm-Leach-Bliley Act). However, both the data and theory²⁸ suggest that the vast majority of banks hold a buffer of equity above the regulatory minimum so as to lower the risk of an adverse shock leading to capital inadequacy. Since the model abstracts from this buffer stock behavior by assuming away any shocks, one would want to include this buffer in the measurement of γ as it is due to the capital adequacy regulation in the first place.²⁹ There is little reason to expect the buffer itself would change dramatically in response to a change in the regulatory minimum capital ratio. In any case, as mentioned, the point is not quantitatively very important. For example, as we change the measure of γ from an unreasonably low value, say, 0.04 to an unreasonably high value, say 0.15, holding constant the other measurements, the estimated marginal welfare cost increases only by a factor 1.13 ($= (1 - 0.15)^{-1} / (1 - 0.04)^{-1}$).

Finally, a measure of the required return on (bank) equity is needed. Since the model abstracts from aggregate risk, a risk-adjusted measure is needed. To avoid the difficulties inherent in measuring the (ex ante) risk premium on regular

²⁶ All variables are nominal. While the model is real, using nominal data is fine, as long as all data are nominal, because the formula for the welfare cost in (26) contains only ratios of quantities and spreads of returns.

²⁷ This may seem incorrect if the capital requirement is not binding. However, if that is the case, the model implies that $R^E = R^D$, so the welfare cost is zero regardless of how γ is measured.

²⁸ See Van den Heuvel (2004) for a quantitative model.

²⁹ In addition these ratios apply to risk weighted assets and off-balance sheet items, considerations from which the model also abstracts.

equity,³⁰ the measure I use is the average return on subordinated bank debt. The reason for this choice is that (a) subordinated debt counts towards regulatory equity capital, albeit within certain limits, and (b) defaults on this type of debt have historically been very rare, so the debt is not very risky. As a measure for $(R^E - 1)$, the net return on subordinated debt is calculated as Interest on Subordinated Notes and Debentures divided by Subordinated Notes and Debentures.³¹

The limits on the use of subordinated debt for regulatory purposes imply that this is a conservative measure for the risk-adjusted required return on bank equity. First, because it is regarded as an inferior form of equity, subordinated debt can count only towards tier 2 capital. Second, and more importantly, the amount of subordinated debt is limited to 50 percent of the bank's tier 1 capital. What this means is that if the tier 1 capital ratio is close to binding, subordinated debt can count for at most approximately 25 percent of total capital. Since banks may use subordinated debt to meet their capital requirements only up to these limits (and they do not have to use it), it is possible that for many banks the required return on subordinated debt is lower than the risk-adjusted return on regular equity.

To measure the welfare cost using the derived formula I compute long run averages for the deposit consumption ratio, for the spread between the return on subordinated debt and deposits, and for the capital asset ratio. The Basle Accord and the FDIC Improvement Act enacting it were not fully implemented until January 1, 1993. For this reason, the sample period is set at 1993-2002. Using 1986-2002 as a longer sample yields very similar results. (1986 is the first year that regulation Q is fully phased out.)

For 1993-2002 the mean deposit to consumption ratio is 0.61, the average net returns on deposits and subordinated debts are, respectively, 3.08% and 6.26%, so the average spread is 3.18%, and the mean capital asset ratio is 0.096.

³⁰ For example, the historical average excess return on bank equity would imply a high premium, but does this equal the *ex ante* expected premium? In addition, depending on what interest rate is used to measure the excess return on equity, this approach runs the risk of contaminating the measured risk premium with a liquidity premium, which one would definitely want to avoid in the present context. If on the other hand one takes a model based measure of the *ex ante* risk premium based on 'reasonable' standard preferences, one would likely get a much lower measure. (This is the well known equity premium puzzle.)

³¹ A part of what is counted in the HSOB as Subordinated Debt and Debentures does not qualify as regulatory tier 2 capital (e.g. if the maturity is too short). However, cross-checking with the Reports on Condition and Income ('call reports') item RCFD5610 indicates that the difference is minimal after 1992.

Hence, applying (26), a first order approximation to the welfare cost of raising the capital requirement by $\Delta\gamma$ is:

$$\nu(\Delta\gamma) = 0.61 \times 0.0318 \times (1 - 0.096)^{-1} \Delta\gamma = 0.0216 \Delta\gamma$$

To interpret this number, suppose that new regulation increases γ by 0.1 (approximately a doubling the effective capital requirement, from almost 0.1 to 0.2) without any change in bank supervision. The resulting welfare loss is equivalent to a permanent loss in consumption of

$$\nu(0.1) = 0.0216 \times 0.1 \times 100\% = 0.216\%.$$

This is not, in my view, a trivial welfare cost. Some well-known estimates on the welfare costs of business cycles or the welfare gains of implementing the optimal monetary policy rule (taking as given average inflation) are much smaller.

Here is another way to interpret this number. Consider lowering the effective capital requirement by 1 percentage point (to 0.086). And suppose regulators can keep the probability of bank failure as low as it is today despite this change by spending more on bank supervision and imposing higher compliance costs. If the total cost of keeping the probability of bank failure the same is less than $\nu(0.01)c_{2002} = 0.0216 \times 0.01 \times 7376 = 1.6$ billion \$ per year, then this ought to be done: lowering the capital requirement would be welfare improving in this way.³² If not, the capital requirement ought to be increased.

It should be pointed out that this estimate is conservative in the sense that, as mentioned, the true spread between the required return on equity and deposits may be higher than the one measured here due to the limits on the use of subordinated debt for regulatory purposes. Section 5 will provide some alternative measures of the spread and associated welfare cost.

A different objection one might have to the above calculation of the welfare cost is that it does not take into account any resource costs that banks incur in servicing deposits or making loans. The former include the costs of ATM networks, part of the cost of maintaining a network of branches, etc. The latter include the costs of screening loan applications, collecting payments, as well as part of the cost of maintaining a branch network. These costs are not trivial. For the period 1993-2002 net noninterest costs of U.S. banks have averaged 1.29% of total assets. The next section will address this concern by incorporating into the

³² Of course, in reality taxation is not lump-sum but usually distortionary, so one would want to make some allowance for that.

model resource costs associated with accepting deposits and/or making loans. Section 5 will use the results of the model to show how this affects the measured welfare cost of the capital requirement.

4. Costly financial intermediation

This section extends the model to allow for resource costs associated with servicing deposits and/or making loans. More precisely, it is now assumed that a bank with D in deposits and L in loans pays a cost $g(D,L)$ to service those financial contracts. The cost enters negatively in the calculation of the bank's net cash flow.

I make the following assumptions on g : $g(D,L)$ is nonnegative, twice continuously differentiable, increasing in its first argument, strictly increasing in its second argument, convex, and homogenous of degree 1, i.e. it exhibits constant returns to scale. Note that linear costs are included as a special case.

For the rest, the model is the same as presented in section 1. It is, however, convenient to impose the following additional assumption on the utility function:

$$\lim_{d \downarrow 0} \frac{u_d(c,d)}{u_c(c,d)} > \frac{g(1-\gamma,1)}{1-\gamma} \text{ for all } c > 0 \quad (28)$$

This ‘weak Inada’ condition³³ is imposed only to streamline the analysis of the equilibrium. If it fails to hold, there is an additional –empirically irrelevant– case to consider in which banks do not exist in equilibrium because the cost of intermediation is too high relative to the marginal value of liquidity, regardless of how scarce liquidity is. If that case applies, the model in any case closely resembles a standard growth model.³⁴ The above assumption is sufficient, but not necessary, to rule out this empirically uninteresting case.

The introduction of the cost of intermediation has a direct effect only on the bank's decision problem. The rest of this section analyses the bank's decisions in the presence of g , and then moves on to describe how the equilibrium changes.

³³ $\lim_{d \downarrow 0} u_d(c,d) = \infty$ is a sufficient condition for the assumption to hold.

³⁴ Obviously, without banks, the welfare cost of increasing the bank capital requirement would be zero in this case. The elements in the formula for the welfare cost, derived below, would be unobservable.

Banks under costly financial intermediation

With costly intermediation, the bank's cash flow is equal to $(R^L + \sigma\varepsilon)L - R^D D - g(D, L)$. With this modification, the bank's problem can still be analyzed in much the same way as in section 1. To avoid repetition this is left for an appendix. Appendix C.1 proves the following results. The bank's choice still involves zero or maximum risk. A sufficient condition for $\sigma = 0$ is:

$$\bar{\sigma} \leq R^L - R^D(1 - \gamma) - g(1 - \gamma, 1) \quad (29)$$

This is also a necessary condition when the capital requirement is binding.

Under the assumption that (29) holds, the analysis of the full shareholder value maximization problem in the appendix shows that

1. If the capital requirement is slack, then

$$R^E = R^D + g_D(D, L)$$

$$R^L = R^E + g_L(D, L)$$

2. If the capital requirement binds, then $E = \gamma L$ and

$$R^L = \gamma R^E + (1 - \gamma)R^D + g(1 - \gamma, 1) \quad (30)$$

In either case, $V^B(E) = E$.

The first case, a nonbinding capital requirement, requires that the bank be indifferent at the margin between financing its assets with deposits or equity and that the bank makes zero profit on making one more loan and financing it with equity (or deposits). The second requirement entails $R^L > R^E$, since $g_L(D, L) > 0$ by assumption. However, if $R^L > R^E$, nonfinancial firms will strictly prefer equity finance to bank loans, as explained in the discussion of firms in section 1. There will be no demand for loans in that case,³⁵ and banks will not exist. Hence, if banks exist in equilibrium, ***with costly financial intermediation, the capital requirement always binds.***

³⁵ Technically, the demand for bank loans would be negative infinity. That is why $R^L \leq R^E$ is a requirement for a finite solution to the firm's problem. See equation (13) in section 1.

With the capital requirement binding, equation (30) holds. Under costless financial intermediation, the rate on bank loans was simply equal to the appropriately weighted average of the returns on equity and deposits. Now, the return on bank loans also includes the term $g(1-\gamma, 1)$ - the resource cost of lending one additional unit and servicing $1-\gamma$ additional units of deposits to finance the loan. As constant returns to scale of g implies that the marginal cost of increasing the scale equals the average cost, the bank still has zero profits: $V^B(E) = E$.

Note that the condition for $\sigma = 0$, stated in (29), is seen to be equivalent to

$$\bar{\sigma} \leq \gamma R^E,$$

the same condition as with $g = 0$ (condition (10)). Since the capital requirement is binding, this condition is necessary as well as sufficient.

Other agents

As mentioned, the cost of intermediation has a direct effect only on the bank's decision problem. Nonetheless, a few comments on the decisions of nonfinancial firms and on the government are in order. For households, the analysis is exactly the same as before.

For nonfinancial firms, it has been shown that when $R_t^L = R_t^E$, firms are indifferent between equity and bank loans. In section 1 it was possible to assume without loss of generality that in that case firms opt for bank loans only. As will be clear shortly, with costly financial intermediation, such an assumption would now no longer be correct, so it is dropped and (14) is replaced by the original optimality condition (13). Finally, the government's balanced budget constraint is now:

$$T_t = T + 0.5 \left(R_t^D D_t - (R_t^L - \sigma_t(1 + 2\xi)) L_t - g(D_t, L_t) \right)^+ \quad (31)$$

General equilibrium with costly intermediation

The definition of equilibrium is the same as in section 1, except that, of course, the bank's maximization problem referred to in the definition is now the

one presented in this section, (14) is dropped, taxes are now set according to (31), and the market clearing condition for the goods market is altered to:

$$F(K_t, 1) - \xi \sigma_t L_t + (1 - \delta) K_t = c_t + K_{t+1} + g(D_t, L_t) + T$$

Again, I focus on the case that (10) holds: $S(T) \leq \gamma R_t^E$, so that $\sigma_t = 0$ and $T_t = T$. Combining the market clearing conditions with equations (1), (2), (30), (11), (12) and (13) it is possible to characterize the equilibrium in terms of a system in (K_t, c_t) with R_t^E and d_t as auxiliary variables:

$$K_t = F(K_{t-1}, 1) + (1 - \delta) K_{t-1} - c_{t-1} - g(1 - \gamma, 1) d_{t-1} / (1 - \gamma) - T \quad (32)$$

$$\beta^{-1}(u_c(c_{t-1}, d_{t-1}) / u_c(c_t, d_t)) = R_t^E \quad (33)$$

$$F_K(K_t, 1) + 1 - \delta = R_t^L = R_t^E - \Delta^+ \quad (34)$$

where $\Delta \equiv (1 - \gamma) \frac{u_d(c_t, (1 - \gamma) K_t)}{u_c(c_t, (1 - \gamma) K_t)} - g(1 - \gamma, 1)$ and d_t is determined as follows:

(a) If $\Delta \geq 0$, firms rely solely on bank loans and

$$d_t = (1 - \gamma) K_t \quad (35)$$

Remark: $R_t^L = R_t^E - \Delta \leq R_t^E$, $L_t = K_t$ and $E_t^F = 0$. Remaining variables are determined through (2) and (11) with $H_t = 1$.

(b) If $\Delta < 0$, firms rely on a mix of equity and bank loans and

$$(1 - \gamma) \frac{u_d(c_t, d_t)}{u_c(c_t, d_t)} = g(1 - \gamma, 1) \quad (36)$$

Remark: $R_t^L = R_t^E$, $L_t = d_t / (1 - \gamma) < K_t$ and $E_t^F = K_t - L_t > 0$. Remaining variables are determined through (2) and (11) with $H_t = 1$. Assumption (28) guarantees that there exists a d_t between 0 and $(1 - \gamma) K_t$ such that (36) holds.

Thus, the spread between the required return on equity and the rate on bank loans is reduced by the cost of financial intermediation. However, if the cost of banking is sufficiently small relative to the value of liquidity services, bank loans are still strictly cheaper than equity finance and the equilibrium is

qualitatively similar to the one with costless intermediation (and a binding capital requirement).

In contrast, if the cost of banking is high, in the sense that $\Delta < 0$, then firms will use both equity and bank finance, in such proportion that, in equilibrium, their costs are exactly equal: $R_t^L = R_t^E$. This is consistent with zero profits in banking only if $(1 - \gamma)$ times the spread between R_t^E and R_t^D exactly compensates for the resource cost of lending one unit and servicing $(1 - \gamma)$ units of deposits, whence equation (36).³⁶ This condition determines both the equilibrium volume of deposits and the right proportion between equity and bank finance, as $L_t = d_t / (1 - \gamma)$ with the binding capital requirement.

It is interesting to note that in case (b) steady state superneutrality of the capital requirement γ holds, in contrast to the result under costless financial intermediation. The reason is that $R_t^L = R_t^E$ in case (b) and the latter is simply β^{-1} in steady state; this pins down the steady state capital stock through (34), which is then no different from the steady state capital stock obtained without any preference for liquidity (i.e. if $u_d \equiv 0$). In case (a), however, it is easy to verify that superneutrality fails for general $u(\cdot)$ as in section 1.

In either case, though, I will show that welfare is affected by the capital requirement. The strategy for doing so is the same as in section 2. I present a constrained social planner's problem that replicates the decentralized equilibrium and then use that equivalence to derive a new formula for the marginal welfare cost of increasing γ .

The social planner's problem with costly intermediation

Define the following constrained social planner's problem:

$$\begin{aligned}
 V_0^{ci}(\theta) = & \max_{\{c_t, d_t, L_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t, d_t) \\
 \text{s.t.} \quad & F(K_t, 1) + (1 - \delta)K_t = c_t + K_{t+1} + g(d_t, L_t) + T \\
 & (1 - \gamma)L_t - d_t \geq 0 \\
 & K_t - L_t \geq 0
 \end{aligned}$$

³⁶ The capital requirement still binds because $R_t^E - R_t^D = g(1 - \gamma, 1) / (1 - \gamma) > g_D(1 - \gamma, 1)$, so the spread exceeds the marginal resource cost of deposits.

Recall that $\theta = (\gamma, T, \delta, \beta, K_0)$. This problem differs from the one in section 2 not only by including the intermediation cost g , but also by allowing for the possibility that $E_t^F = K_t - L_t > 0$.

Appendix C.2 displays the first-order conditions to this problem and uses these to characterize the solution. It also shows that the resulting allocation is identical to the decentralized equilibrium with costly intermediation when $\sigma = 0$ in that equilibrium (i.e. when $S(T) \leq \gamma R_t^E$). In other words, *the constrained social planner's problem replicates the decentralized equilibrium* and, as a consequence, *welfare in the decentralized equilibrium equals $V_0^{ci}(\theta)$* , the value of the objective function to the constrained social planner's problem with costly intermediation.

A formula for the marginal welfare cost for the case of costly intermediation

Again, I exploit the equivalence of the constrained social planner's problem and the decentralized equilibrium to measure the marginal effect on welfare of a change in the capital requirement. Going through the same steps as in section 1, appendix C.3 proves the following proposition:

Proposition 2 *Assume the economy is in steady state in period 0 and (10) holds. Consider permanently increasing γ by $\Delta\gamma$. With costly financial intermediation, a first-order approximation to the resulting welfare loss, expressed as the welfare-equivalent permanent relative loss in consumption, is*

$$V^{ci}(\Delta\gamma) = \frac{d_0}{c_0} (R_0^E - R_0^D - g_D(d_0, L_0)) (1 - \gamma)^{-1} \Delta\gamma \quad (37)$$

This formula differs from the one derived for costless intermediation in that it subtracts the marginal resource cost of servicing deposit contracts from the spread between the returns on equity and deposits.³⁷ The intuition for this change is straightforward: If liquidity creation is costly, then, even in the absence of a binding capital requirement, this creates a spread between the returns on equity and deposits, as banks need to be compensated for this cost. It is only to the extent that the spread exceeds the marginal resource cost of deposits that a scarcity of

³⁷ Note that if $g = 0$, the formula specializes to the one derived in proposition 1 for costless intermediation.

deposits due to the binding capital requirement is revealed. Only then is there a welfare effect at the margin.

It is worth emphasizing that the formula is valid whether the equilibrium is characterized by pure bank finance or by mixed bank and equity finance. How one might measure, or bound, the marginal cost of deposits is addressed in the next section.

5. Measurement of the welfare cost, part 2

Arguably the most conservative way of measuring the new term, $g_D(D, L)$, in the expression for the welfare cost (37) is to calculate upper and lower bounds based only on assumptions already made, namely that the cost function g is nondecreasing and exhibits constant returns to scale. These imply:³⁸

$$0 \leq g_D(D, L) \leq \frac{g(D, L)}{D}$$

Setting $g_D = 0$ yields the same result as with costless financial intermediation. Section 3 found that $\nu(\Delta\gamma) = 0.0216\Delta\gamma$ in this case. To implement the upper bound, g is measured as net noninterest cost (Total Noninterest Expense minus Total Noninterest Income). The average ratio of net noninterest cost to deposits for 1993-2002 is 0.0187 (i.e. 1.87 percent). With this upper bound for g_D we get the following *lower* bound for the welfare cost:

$$\nu^{ci}(\Delta\gamma) \geq 0.61 \times (0.0318 - 0.0187) \times (1 - 0.096)^{-1} \Delta\gamma = 0.0089\Delta\gamma$$

According to this method, the welfare loss of raising the capital requirement by $\Delta\gamma = 0.1$ is then somewhere between 0.089% and 0.216% of consumption (permanently). Naturally, recognizing that financial intermediation is costly leads to a somewhat lower estimate of the marginal welfare cost.

As mentioned, using subordinated debt to measure the required return on equity is likely a conservative way to measure the welfare cost. We can measure the spread in an alternative way by rewriting the bank's zero profit condition (30), as follows:

³⁸ With decreasing returns to scale these bounds would also apply. The assumption that $g_L > 0$ actually implies that the second inequality is strict.

$$R^E - R^D = \gamma^{-1} \left(R^L - R^D - g(1-\gamma, 1) \right)$$

Using the constant returns to scale of g and the result that the capital requirement binds, this yields an alternative, theoretically equivalent, way of measuring the welfare cost:

$$\nu^{ci}(\Delta\gamma) = \frac{d}{c} \left(\frac{1}{\gamma} \left(R^L - R^D - \frac{g(D, L)}{L} \right) - g_D(D, L) \right) (1-\gamma)^{-1} \Delta\gamma \quad (38)$$

To implement this, I use two alternative quantities for measuring loans L : Total Loans and Total Assets, as it is possible to regard securities owned by banks as bank loans in another form. When using Total Loans, the net return on loans ($R^L - 1$) is calculated as (Total Interest Income on Loans minus the Provision for Loan Lease Losses) divided by Total Loans. The Provision for Loan Lease Losses represents “the amount needed to make the allowance for loan and lease losses adequate to absorb expected loan and lease losses, based upon management's evaluation of the bank's current loan and lease portfolio” (HSOB). That is, it captures the decline in the value of loans due to an increase in expected default losses. When using Total Assets, ($R^L - 1$) is computed as (Total Interest Income minus the Provision for Loan Lease Losses plus Securities Gains/Losses) divided by Total Assets. All other variables are measured in the same way as in section 3.

I also consider two sample periods for computing the historical averages: the post-Basel/FDICIA period 1993-2002, as before, and 1986-2002. While the second sample periods includes the regulatory changes associated with Basel and the FDICIA (which increased the effective capital requirement somewhat), it has the advantage that it is longer. The reason for not going back further in time, is that regulation Q, which placed some restrictions on banks' setting of deposit rates, was not fully phased out until January 1, 1986. Incidentally, in section 3, I chose not to use the 86-02 sample with subordinated debt, because before the Basel Accord its use in satisfying capital adequacy rules was rather limited.³⁹

The returns on loans and assets are likely to contain a risk premium, albeit a much smaller one than for equity, so using them without risk adjustment may well result in an upwardly biased measurement of the welfare cost. I will first

³⁹ Nonetheless, as can be seen in table 1, the results are quite similar when using 86-02 instead of 93-02 with subordinated debt.

TABLE 1. WELFARE COST OF INCREASING THE CAPITAL REQUIREMENT BY 0.1 ($v^{ci}(0.1)$ in percent).

Sample:	1986 - 2002		1993 – 2002	
	$g_D = 0$	$g_D = g/D$	$g_D = 0$	$g_D = g/D$
Subordinated Debt	0.21	0.06	0.22	0.09
Total Assets	1.05	0.91	1.19	1.07
<i>Risk adjusted</i>	0.82	0.67	1.16	1.03
Total Loans	1.32	1.17	1.56	1.43
<i>Risk adjusted</i>	0.96	0.82	1.46	1.33

Notes: first order approximation to the welfare loss associated with $\Delta\gamma = 0.1$, expressed as the welfare equivalent percent permanent loss in consumption. The first row implements (37), all other rows are based on (38).

present results using unadjusted returns and then construct risk-adjusted measures and use those.

For the longer sample period (1986-2002), the unadjusted return on Total Assets averages 6.75%. After deducting noninterest cost (1.51% of total assets) this exceeds the return on deposits (3.96%) by 128 basis points. As a result, using (38), the first-order approximation to welfare cost of raising the capital requirement by $\Delta\gamma = 0.1$ is found to be equivalent to a permanent cut in consumption of between 0.91% and 1.05%, depending on the cost share of deposits. This is considerably higher, by about a factor 5, than the previous estimates, which used subordinated debt.

For comparability, table 1 displays the results of the various strategies for measuring the welfare cost. The numbers are the permanent loss in consumption, in percent, that is welfare equivalent to $\Delta\gamma = 0.1$, without any change in banking supervision. As can be seen in the fourth row of the table, using Total Loans rather than Total Assets results in an estimate of the welfare cost that is even a little higher, ranging from 1.17% to 1.32%. The reason is that the return on total loans is, at 8.07%, exceeds the return on total assets. Using the 1993-2002 sample also results in slightly higher measurements of the cost in all cases.

While the estimates using subordinated debt should be considered conservative, there is, as mentioned, a concern that these new results overstate the

welfare cost, because in reality the returns on total assets or loans likely contain a nontrivial risk premium, which is absent in the model. To examine to what extent this accounts for the higher estimates, I perform a crude, back-of-the-envelope risk adjustment, as follows.

The historical standard deviation of the spread between loans and deposits, net of noninterest cost, i.e. of $R_t^L - R_t^D - g(D_t, L_t) / L_t$, is 0.57% for total assets, or 0.86% for total loans.⁴⁰ Treating R_t^D as a risk free rate, the resulting Sharpe ratio of $R_t^L - g_t / L_t$ is about 2 in each case. Compared to the Sharpe ratio of the US stock market, approximately 0.5 for the S&P500, this seems very high. To view the average excess return $R_t^L - g_t / L_t - R_t^D$ as purely a risk premium, one must believe that bank loans are about 4 times as risky per unit standard deviation as the stock market. That does not seem very plausible to me. To the extent that one regards the equity premium as a puzzle, it is then an even greater puzzle to explain this ‘banking premium’, if it is purely a reward for risk.

Thus, borrowing by accepting (FDIC-insured) deposits and making bank loans with the proceeds is a strategy with very favorable risk-return properties. However, it is not actually possible to execute this strategy exactly if the capital requirement binds – some of the loans must be financed with equity. And this is precisely the point: to the extent that the capital requirement binds, the high Sharpe ratio of bank loans relative to deposits is not such a puzzle, because the model predicts a positive ‘banking premium’ even in the absence of risk.

This discussion suggests a very simple way of risk-adjusting the return on loans. Assume the market price of risk equals the Sharpe ratio of the stock market, roughly 0.5. In addition, assume that all the variation in the excess return on loans is risk.⁴¹ Under these assumptions, the risk premium in $R_t^L - g_t / L_t - R_t^D$ is $0.5 \times 0.57\% = 0.29\%$ for total assets, or $0.5 \times 0.86\% = 0.43\%$ for total loans. Deducting this risk premium from the spread lowers the measured welfare cost moderately. For the longer sample, the risk-adjusted estimates for the welfare cost of an increase in the capital requirement by 0.1 range from 0.67 to 0.96 percent (see table 1).

These numbers are similar to the estimated welfare cost of permanently increasing inflation from zero to 10 percent, as measured by Lucas (2000). They are also close to the low end of the range of estimates of the welfare gain of

⁴⁰ These numbers are for 1986-2002; for 1993-2002 they are 0.10% and 0.28%, respectively.

⁴¹ For example, the CAPM holds and both the stock market and bank loans have a beta equal to one.

eliminating capital income taxation, when taking into account the welfare effects of the transition to a new steady state (see, e.g., Lucas 1990).

Admittedly, the risk adjustment is outside the model and the calculation is rather crude. Nonetheless, it suggests that the welfare cost of capital requirements may well be underestimated considerably by using subordinated debt as a proxy for the required return on equity, though the difference is almost certainly less than an order of magnitude. Despite the limited remaining uncertainty, taken together, these results suggest a fairly large welfare cost of bank capital adequacy regulation.

6. Are capital requirements too high or too low?

What does the sizable welfare cost imply for optimal bank regulation? It is important to realize that the welfare cost measured here is a *gross* cost.⁴² It needs to be compared with the benefit of capital requirements in reducing the cost of bank supervision and compliance. As discussed in section 2, with a higher capital requirement it is easier for bank supervisors to prevent excessive risk taking by banks, so that spending on bank supervision, T , can be lowered. It is conceivable that the resulting marginal welfare benefit (derived in equation (27)) is as large as the marginal welfare cost, in which case current regulations are in fact optimal. This section is an attempt to measure this marginal welfare benefit in order to find out whether current capital requirements are too high, too low or about right.

Before moving on to the calculations involved in this exercise, it is worth emphasizing that if one believes that deposit insurance does not create a moral hazard problem (perhaps because bank risk is perfectly and costlessly observable), or that deposit insurance is undesirable to start with (and there are no other externalities associated with bank failures), then capital requirements have no benefits at all, and the welfare cost measured above is a net cost as well as a gross cost.

⁴² It has been argued that a similar point applies to standard estimates of the welfare cost of inflation: increasing inflation results in more seigniorage, so that taxes can be lowered while keeping government revenue constant. If taxes are distortionary, then the *net* welfare cost of higher inflation is lower. However, Lucas (2000) finds that the resulting adjustment is small. The measures of the welfare gain of reducing capital income taxation reported in Lucas (1990) are truly *net* gains.

This paper has instead adopted the view that there exists a rationale for a capital requirement. Its marginal welfare benefit, in steady state, is given by the right hand side of equation (27) (in units of the good per period):

$$mwb = -\frac{dT}{d\gamma} \bigg|_{S(T)=\gamma\beta^{-1}} = \frac{-1}{\beta S'(T)} \quad (39)$$

where it is recalled that T is spending of bank supervision and S is the ‘supervision technology’ mapping T into $\bar{\sigma}$, the maximum undetectable level of risk taking, $S' \leq 0$.

As this expression depends on the particular form of the incentive compatibility constraint, $S(T) \leq \gamma R^E$, the reader may be concerned that the form of this constraint in turn depends on the very special assumption made with regard to the distribution of the bank-specific shock for excessive risk, ε , given in equation (3). This is not the case, however: Appendix D shows that, with an appropriate normalization, this incentive compatibility constraint is valid for any distribution of ε that has a nonpositive mean and bounded support $[\underline{\varepsilon}, \bar{\varepsilon}]$, with $\underline{\varepsilon} < 0 < \bar{\varepsilon}$.

To evaluate the marginal welfare benefit, it is clear that a measure of the marginal effectiveness of bank supervision spending is required. This is no easy task. Stronger assumptions are needed than the ones that were required for the measurements of the welfare cost. Conclusions based on the calculations below are therefore best regarded as somewhat tentative.

To measure the marginal effectiveness of bank supervision, I make two assumptions. First,

(i) I use the following log-linear approximation for $S(T)$:

$$\ln S(T) = \ln S(0) - \alpha T, \quad \alpha \geq 0 \quad (40)$$

This approximation has the virtue of satisfying all the assumptions made on S , in particular that it is positive for all T . It also has the appealing feature that $S'' \geq 0$, implying a diminishing marginal effectiveness of bank supervision spending.⁴³

⁴³ It seems natural to impose that $S(0) = \sigma_{RF}$, since in the absence of bank supervision, presumably nothing would prevent banks from making all their loans to risky firms, if they wished to do so, which would imply $\sigma_t = \sigma_{RF}$. But nothing in the calibration below depends on having $S(0) = \sigma_{RF}$, so I do not impose it.

Rewriting (40) as $S(T) = S(0)e^{-\alpha T}$, we have $S'(T) = -\alpha S(T)$. Inserting this into (39) yields

$$mwb = \frac{1}{\beta \alpha S(T)}$$

In light of the very low rate of bank failures in the U.S. since the implementation of the Basel Accord,⁴⁴ I regard the current U.S. regulatory regime as largely successful in preventing excessive risk taking. That is, it satisfies the incentive compatibility constraint:⁴⁵ $S(T) \leq \gamma \beta^{-1}$.

If the inequality is strict, then the *marginal* welfare benefit of the capital requirement in fact equals zero: at the margin γ could be lowered without any increase in supervision spending. If the marginal welfare cost is any strictly positive number, then welfare can be improved by lowering the capital requirement. (This is of course a local calculus of variation argument. More can be said about this case and I will return to it below.)

If, on the other hand, $S(T) = \gamma \beta^{-1}$, we have

$$mwb = \frac{1}{\alpha \gamma} \tag{41}$$

To measure α I make the following second assumption:

(ii) I assume that in the absence of bank supervision there is a positive probability that a bank that engages in maximum excessive risk taking will realize a (near) total loss for the deposit insurance fund as a fraction of the bank's deposit liabilities. Or, equivalently, in the absence of bank supervision, there is a positive probability, however small, that the realized gross return to lending is (close to) zero if the bank has taken on maximum excessive risk.

To see the equivalence, note that the loss to the deposit insurance fund from a single bank evaluated at $\underline{\varepsilon}$, the lowest level of the support of ε , and $\sigma = S(0)$, maximum risk taking in the absence of supervision, is:

⁴⁴ During 1993-2002 about 0.1 percent of all FDIC insured commercial banks failed on average per year.

⁴⁵ I continue to assume that the U.S. economy is in steady state, so $R^E = \beta^{-1}$.

$$R^D D - (R^L + S(0)\underline{\varepsilon})L + g(D, L)$$

According to assumption (ii), this should be approximately equal to $R^D D$. This requires

$$R^L - g(D, L)/L + S(0)\underline{\varepsilon} \approx 0$$

In other words, with zero supervision and maximum risk taking, there is some (arbitrarily small, but strictly positive) probability of realizing a near zero gross return on the bank's loan portfolio, net of noninterest cost. Given that the probability can be arbitrarily small, and given the opportunities for banks that currently exists for taking on more risk, e.g. through derivatives, this does not strike me as a very strong assumption.

This condition can be used to infer $S(0)$, which will be useful. With the example distribution in (3), $\underline{\varepsilon} = -(1 + 2\xi)$, where ξ is the absolute value of the negative mean of ε . Thus, ξ is the direct loss in net present value due to excessive risk per unit of bank risk, σ . In modern financial markets, there exist excellent opportunities for banks to take on more risk, not only by making risky loans, but also through the trading book and through various off-balance sheet activities. It is thus hard to believe that ξ is a large number in reality. I therefore assume that $\xi \approx 0$. For the example distribution, this implies $\underline{\varepsilon} \approx -1$. The same is true for the general distribution analyzed in Appendix D: as shown there, with the appropriate normalization that yields the desired form of the incentive compatibility constraint, $\underline{\varepsilon} \approx -1$ when $\xi \approx 0$. Using this,

$$S(0) \approx R^L - g(D, L)/L$$

The right hand side is larger than 1 (based on the data) and less than β^{-1} (the model's steady state return on equity). The difference is trivial for our purposes, so I simply set $S(0) = 1$. Thus, $S(T) = e^{-\alpha T} \geq \gamma\beta^{-1}$ (by incentive compatibility), so

$$\alpha \leq -\ln(\gamma\beta^{-1})/T,$$

with equality if $S(T) = \gamma\beta^{-1}$.

Total spending on bank supervision by the OCC, the FDIC, the Federal Reserve System, and state agencies was \$1.4 billion in 1999 (Hawke (2000)).

Although some of this spending is related to regulations that are only tangentially related to promoting the ‘safety and soundness’ of the banking system,⁴⁶ I conservatively set $T = \$1.4$ billion per year. Using $\gamma = 0.1$, and $\beta^{-1} = 1.06$ we have $\alpha \leq 1.6$.

If the incentive compatibility constraint currently binds, i.e. $S(T) = \gamma\beta^{-1}$, then $\alpha = 1.6$ and (41) can be applied to calculate the marginal welfare benefit:

$$mwb = 1/\alpha\gamma = 6.2 \text{ billion \$ per year.}$$

This number may have an upward bias because of the inclusion of all bank supervision spending. On the other hand, it does not reflect the banks’ cost of compliance with the supervisory process. If we factor in compliance cost and assume it is *twice* bank supervision spending, we have $T = \$4.2$ billion, $\alpha \leq 0.53$, and, if the incentive compatibility constraint binds, the inferred marginal welfare benefit rises to \$18.7 billion per year.

As mentioned, this marginal welfare benefit must be compared to the marginal welfare cost of capital requirements, in units of the good per year, $\nu^{ci}(1)c$ (see equation (27)). Using the estimates for the welfare cost based on subordinated debt, as well as the risk-adjusted ones using total assets and loans, the marginal welfare cost $\nu^{ci}(1)c$ ranges from \$57 billion to \$603 billion per year, with a mean of \$375 billion.⁴⁷ Even using the lowest number, which is based on subordinated debt and attributes *all* net noninterest cost to servicing deposits, it appears that the marginal welfare cost substantially exceeds the marginal welfare benefit. The conclusion is that capital requirements are currently set *too high*: welfare can be raised by lowering them and the welfare gains of doing so are potentially large.

It is true that the calculation of the marginal welfare benefit relies on a functional form approximation and some additional assumptions. It is certainly possible to obtain other and perhaps better measures of this quantity. Nonetheless, I would be surprised if, based on 1.4 \$ billion in supervision spending, one would find a number in the hundred billion-plus range. To regard the present regulatory

⁴⁶ To name just a few examples from among many: examining for compliance with the Fair Lending Act, Right to Financial Privacy Act, National Flood Insurance Act, etc.

⁴⁷ See table 1. This uses preferred sample periods (1993-2002 for sub-debt and 1986-2002 for total assets and total loans) and 1999 consumption, which is \$6283 billion, for comparability with the marginal welfare benefit. For example, the lower bound of the range is based on sub-debt, $g_D = g/D$ and sample 93-02: $\nu^{ci}(0.1) = 0.09\%$, so $\nu^{ci}(1)c = (0.09/100\%) \times (1/0.1) \times 6283 = 57$.

environment as optimal, one has to believe that there is something magical about the current 10% capital ratio; that a slight decrease in this ratio will lead to a sudden and large increase in the number of bank failures from the near zero rate today, and that this increase is not preventable by, say, a doubling or tripling of supervision spending.

What should happen to bank supervision spending? This is not clear. If the incentive compatibility constraint currently binds, i.e. if $S(T) = \gamma\beta^{-1}$, then the conclusion that γ should be lowered implies that supervision spending, T , should be increased to compensate ($S' < 0$). However, if the constraint does not bind, such a conclusion is not warranted. In that case, it may well be optimal to lower both the capital requirement and bank supervision cost.

7. Conclusion

This paper has developed a framework for measuring the welfare cost of bank capital requirements. Such requirements can be socially costly because they reduce banks' ability to create liquidity in equilibrium. Using U.S. data, I have measured this cost in a variety of ways. According to the most conservative estimates, the welfare cost of a 10 percentage point increase in capital requirements is equivalent to a permanent loss in consumption of 0.1 to 0.2 percent. The other measurements find a cost equal to slightly less than 1 percent of consumption. This is a fairly large welfare cost.

Moreover, this cost is much larger than the measured benefit of capital requirements in reducing the cost of bank supervision, which is the other tool regulators possess to limit the moral hazard problem associated with deposit insurance. Though the measurement of the welfare benefit required stronger assumptions, it thus appears that capital requirements are currently too high.

Regulators face an important trade-off between, on the one hand, keeping the effective capital requirement ratio as low as possible and, on the other hand, limiting the supervision and compliance cost associated with capital adequacy regulation, all the while keeping the probability of bank failure acceptably low. It is thus not obvious that the current trend towards a more complex regulatory regime is outrageously wrong. But the stated goal of keeping capital ratios at about the same level for the average bank is not justified.

APPENDIX A. RISKY FIRMS

Since there is no asymmetric information between the bank and the firm, the optimal financial contract will have the repayment depend on the realization of the shock ε . Let $\tilde{R}^L(\varepsilon)$ denote the contractual loan repayment rate as a function of the shock. Profits are of a risky firm are:

$$\tilde{\pi}^F = F(K, H) + \sigma_{RF} \varepsilon K + (1 - \delta)K - wH - \tilde{R}^L(\varepsilon)K$$

(It is straightforward to verify that no household is willing to provide the risky firm with equity.) For ease of exposition define

$$f(K) \equiv \max_H F(K, H) + (1 - \delta)K - wH$$

One of the results in the main text is that, in an equilibrium in which riskless firms exist, they have zero profits and indeterminate scale, so that

$$f(K) = R^L K$$

where R^L is the equilibrium riskless loan rate. Hence, given an optimal choice for H , profits of the risky firm equal

$$\tilde{\pi}^F = R^L K + \sigma_{RF} \varepsilon K - \tilde{R}^L(\varepsilon)K$$

Limited liability of the shareholders implies $\tilde{\pi}^F \geq 0$ in each state. Hence,

$$\tilde{R}^L(\varepsilon) \leq R^L + \sigma_{RF} \varepsilon$$

The right hand side is the most the bank can charge in each state without violating limited liability. Suppose the loan rate equals this upper bound in each state. Then, from assumption (3), it follows that

$$\mathbb{E}_\varepsilon[\tilde{R}^L(\varepsilon)] = R^L - \xi \sigma_{RF} < R^L$$

Since this is still a worse expected return than for a nonrisky loan, the risky firm cannot hope to get better terms, so that, in fact, $\tilde{R}^L(\varepsilon) = R^L + \sigma_{RF} \varepsilon$, if any lending to risky firms occurs.

With this loan contract, the risky firm has zero (expected) profits, so its participation constraint is satisfied. As mentioned in the main text, this implies that a bank can create a portfolio of riskiness σ by directing a fraction σ / σ_{RF} of its lending to risky firms with perfectly correlated shocks (e.g. one risky firm).

Finally, it is easily verified that labor demand of a risky firm satisfies the same first-order condition (11) as for a nonrisky firm.

APPENDIX B. PROOF OF FAILURE OF SUPERNEUTRALITY

First, I state explicitly a convenient assumption that bounds taxes from above in order to guarantee existence of a steady state equilibrium with positive consumption for the assumed utility function (see footnote 21). Define \hat{K} by writing

$$F_K(\hat{K}, 1) + 1 - \delta \equiv \beta^{-1}. \quad (42)$$

Note that \hat{K} exists and is unique. \hat{K} is the steady state level of the capital stock under the assumption $a = 0$ (i.e. $u(c, d) = \tilde{u}(c)$, so that the model is not materially different from the standard growth model). I assume that

$$T < F_H(\hat{K}, 1) \quad (43)$$

That is, taxes are lower than aggregate wage income at \hat{K} .⁴⁸

With $a > 0$, $u_d(c, d) > 0$ everywhere, so the capital stock is determined by the system of equations (16)-(18) and (20), which for a steady state and the assumed functional form of $u(\cdot)$ simplifies to

$$\begin{aligned} F_K(K^*, 1) + 1 - \delta &= \beta^{-1} - (1 - \gamma)a \left(\frac{c^*}{(1 - \gamma)K^*} \right)^{1/\eta} \text{ and} \\ F(K^*, 1) + \delta K^* &= c^* + T \end{aligned}$$

where starred variables denote steady state levels. Equivalently,

$$F_K(K^*, 1) + 1 - \delta = \beta^{-1} - a(1 - \gamma)^{(\eta-1)/\eta} \Psi(K^*)^{1/\eta} \quad (44)$$

with

$$\Psi(K) \equiv (F(K, 1) - T) / K - \delta$$

Since $c^* / K^* = \Psi(K^*)$, $c^* \geq 0 \Leftrightarrow \Psi(K^*) \geq 0 \Leftrightarrow K^* \in [\underline{K}, \bar{K}]$, where \underline{K}, \bar{K} are the two solutions to $F(K, 1) - \delta K = T$.⁴⁹

Using assumption (43), the definition of \hat{K} (42), and Euler's theorem, it is straightforward to show that $\Psi(\hat{K}) > 0$, so that $\hat{K} \in (\underline{K}, \bar{K})$. Moreover, again using Euler's theorem,

$$\Psi'(K) = (T - F_H(K, 1)) / K^2$$

⁴⁸ Recall that $H = 1$ in equilibrium. The assumption is sufficient but not necessary.

⁴⁹ The two roots exist since $\Psi(\hat{K}) > 0$, as explained in the next paragraph, and because F satisfies the Inada conditions.

Define \tilde{K} by writing $F_H(\tilde{K}, 1) \equiv T$. Note that \tilde{K} exists and is unique, $\tilde{K} < \hat{K}$ by (43) and $\Psi'(K) > (<)0$ iff $K < (>)\tilde{K}$. The last fact also implies that $\tilde{K} > \underline{K}$.

Since $\tilde{K} < \hat{K} < \bar{K}$, as we let K increase from \tilde{K} to \bar{K} , the left-hand side of (44), $F_K(K, 1) + 1 - \delta$, drops, continuously and monotonically, from a value strictly greater than β^{-1} to a value strictly less than β^{-1} .⁵⁰ Again, as we let K increase from \tilde{K} to \bar{K} , the right-hand side of (44), $\beta^{-1} - (1-\gamma)a\Psi(K)^{1/\eta}$, rises, continuously and monotonically, from a value strictly less than β^{-1} to exactly β^{-1} (as $\Psi(\bar{K}) \equiv 0$). Hence, there is exactly one K^* in $[\tilde{K}, \bar{K}]$ satisfying (44). By a similar argument it is easy to show that there is no K^* in $[\underline{K}, \tilde{K}]$ satisfying (44). Hence, *there exists a unique steady state level of the capital stock*. It is interesting to note that its marginal product is less than β^{-1} , so K^* exceeds \hat{K} , the steady state level of capital without liquidity preference.

Total differentiation of equation (44) with respect to K^* and γ yields:

$$\frac{dK^*}{d\gamma} = \frac{(a/\eta)(\eta-1)(1-\gamma)^{-1/\eta}\Psi(K^*)^{1/\eta}}{F_{KK}(K^*, 1) + (a/\eta)(1-\gamma)^{(\eta-1)/\eta}\Psi(K^*)^{(1-\eta)/\eta}\Psi'(K^*)}$$

$\tilde{K} < K^* < \bar{K}$ implies that $\Psi(K^*) > 0$ and $\Psi'(K^*) < 0$. Since, in addition, $F_{KK}(K^*, 1) < 0$,

$$\text{sign}\left(dK^*/d\gamma\right) = \text{sign}(1-\eta).$$

QED.

APPENDIX C. COSTLY FINANCIAL INTERMEDIATION

1. The bank's problem. The value of the bank to its shareholders right after the bank has raised E in equity at the beginning of the period is now:

$$\begin{aligned} V^B(E) &= \max_{L, D, \sigma} \mathbb{E}_\varepsilon \left[\left((R^L + \sigma\varepsilon)L - R^D D - g(D, L) \right)^+ \right] / R^E \\ \text{s.t. } &L = E + D \\ &E \geq \gamma L \\ &\sigma \in [0, \bar{\sigma}] \end{aligned} \tag{45}$$

⁵⁰ In fact, it is possible to show that $F_K(\bar{K}, 1) + 1 - \delta < 1$ by using the definition of \bar{K} and assumption (43).

The only difference with (4) is the presence of the resource cost $g(D, L)$.

First, a similar argument as in section 1 can be used to characterize the choice of σ conditional on L and D . Expected dividends are

$$\begin{aligned} \mathbb{E}_\varepsilon & \left[\left((R^L + \sigma\varepsilon)L - R^D D - g(D, L) \right)^+ \right] \\ &= \begin{cases} (R^L - \sigma\xi)L - R^D D - g(D, L) & \text{if } (R^L - \sigma(1+2\xi))L - R^D D - g(D, L) \geq 0 \\ 0.5((R^L + \sigma)L - R^D D - g(D, L)) & \text{otherwise} \end{cases} \end{aligned}$$

As before, expected dividends are a piecewise linear, convex function of σ , so the optimal choice of riskiness is at a boundary of the feasible set $[0, \bar{\sigma}]$. By evaluating expected dividends under $\sigma = 0$ versus $\sigma = \bar{\sigma}$, and using the constant returns to scale of g , it is easy to verify that

$$\begin{aligned} \sigma &= 0 \text{ iff } \bar{\sigma} \leq R^L - R^D(D/L) - g(D/L, 1) \\ \sigma &= \bar{\sigma} \text{ otherwise} \end{aligned}$$

Again, for convenience it is assumed that the bank chooses $\sigma = 0$ when bank is indifferent between the two choices at $\bar{\sigma} = R^L - R^D(D/L) - g(D/L, 1)$. Because $E = L - D \geq \gamma L$, a sufficient condition for $\sigma = 0$ is:

$$\bar{\sigma} \leq R^L - R^D(1 - \gamma) - g(1 - \gamma, 1) \quad (29)$$

This is also a necessary condition if $E = \gamma L$, i.e. if the capital requirement is binding. If (29) holds, the bank's sub-problem in (45) simplifies to:

$$\begin{aligned} V^B(E) &= \max_L \left(R^L L - R^D(L - E) - g(L - E, L) \right) / R^E \\ \text{s.t. } & E - \gamma L \geq 0 \end{aligned}$$

where the balance sheet identity, $D = L - E$, has been substituted into the objective function. While this sub-problem is straightforward to solve, it economizes on algebra to characterize the solution as part of an analysis of the bank's full problem, which includes the choice on how much equity to raise. In choosing E , the bank maximizes its pre-issue value, $V^B(E) - E$:

$$\begin{aligned} \pi^B &= \max_E \left(V^B(E) - E \right) \\ &= \max_{E, L} \left(R^L L - R^D(L - E) - g(L - E, L) \right) / R^E - E \\ \text{s.t. } & E - \gamma L \geq 0 \end{aligned}$$

The first-order conditions are:

$$(L) \quad R^L - R^D - g_D(D, L) - g_L(D, L) = \gamma R^E \chi$$

$$(E) \quad R^D + g_D(D, L) = (1 - \chi)R^E$$

where χ is the Kuhn-Tucker multiplier associated with the capital requirement: $\chi \geq 0$ and $\chi(E - \gamma L) = 0$. There are two cases to consider:

If $\chi = 0$, i.e. if the capital requirement is slack, the first-order conditions specialize to

$$R^E = R^D + g_D(D, L)$$

$$R^L = R^E + g_L(D, L)$$

If $\chi > 0$, i.e. if the capital requirement binds, $E = \gamma L$ and the first-order conditions yield

$$R^L = \gamma R^E + (1 - \gamma)R^D + (1 - \gamma)g_D(D, L) + g_L(D, L)$$

Since $D = (1 - \gamma)L$ in this case, and using the fact that the partial derivatives of g are homogenous of degree zero as well as Euler's theorem, this is easily simplified to

$$R^L = \gamma R^E + (1 - \gamma)R^D + g(1 - \gamma, 1)$$

2. Equivalence of the constrained social planner's problem. The Lagrangian and the first-order conditions to this problem are:

$$\ell_0^{ci}(\theta) = \max_{\{c_t, d_t, L_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \{u(c_t, d_t) + \lambda_t^{sp} [F(K_t, 1) + (1 - \delta)K_t - c_t - g(d_t, L_t) - K_{t+1} - T] \\ + \chi_t^{sp} [(1 - \gamma)L_t - d_t] + \mu_t^{sp} [K_t - L_t]\}$$

$$(c) \quad u_c(c_t, d_t) = \lambda_t^{sp}$$

$$(d) \quad u_d(c_t, d_t) = \lambda_t^{sp} g_D(d_t, L_t) + \chi_t^{sp}$$

$$(L) \quad (1 - \gamma)\chi_t^{sp} = \lambda_t^{sp} g_L(d_t, L_t) + \mu_t^{sp}$$

$$(K) \quad \lambda_t^{sp} [F_K(K_t, 1) + 1 - \delta] - \beta^{-1} \lambda_{t-1}^{sp} + \mu_t^{sp} = 0$$

with $\chi_t^{sp} \geq 0$, $\chi_t^{sp} [(1 - \gamma)K_t - d_t] = 0$, $\mu_t^{sp} \geq 0$ and $\mu_t^{sp} [K_t - L_t] = 0$. Since $g_L > 0$, the first order conditions with respect to consumption and loans imply $\chi_t^{sp} > 0$, so that the 'capital requirement' binds: $d_t = (1 - \gamma)L_t$. There are thus only two cases to consider:

(a). If $\mu_t^{sp} > 0$, then $K_t = L_t$ and, since $\chi_t^{sp} > 0$, $d_t = (1-\gamma)L_t = (1-\gamma)K_t$. Rewriting the first-order condition with respect to K ,

$$\begin{aligned} F_K(K_t, 1) + 1 - \delta &= \beta^{-1} \lambda_{t-1}^{sp} / \lambda_t^{sp} - \mu_t^{sp} / \lambda_t^{sp} \\ &= \frac{\beta^{-1} u_c(c_{t-1}, d_{t-1})}{u_c(c_t, d_t)} - \left\{ (1-\gamma) \frac{u_d(c_t, d_t)}{u_c(c_t, d_t)} - g(1-\gamma, 1) \right\} \end{aligned}$$

where the last equality follows the first order conditions with respect to d and L and the homogeneity of g . Since the term in curly brackets equals $\mu_t^{sp} / \lambda_t^{sp}$, it must be strictly positive. As $d_t = (1-\gamma)K_t$ here, this case thus requires that $\Delta > 0$. (Recall that $\Delta \equiv (1-\gamma)u_d(c_t, (1-\gamma)K_t)/u_c(c_t, (1-\gamma)K_t) - g(1-\gamma, 1)$.)

(b). If $\mu_t^{sp} = 0$, the first order conditions yield:

$$\begin{aligned} F_K(K_t, 1) + 1 - \delta &= \beta^{-1} u_c(c_{t-1}, d_{t-1}) / u_c(c_t, d_t) \\ u_d(c_t, d_t) &= u_c(c_t, d_t) g(1-\gamma, 1) / (1-\gamma) \end{aligned}$$

The second equation is obtained by combining the first order conditions with respect to d and L . As case (a) requires $\Delta > 0$, case (b) must apply if $\Delta \leq 0$. If $\Delta \leq 0$, then by assumption (28) there exists a positive $d_t \leq (1-\gamma)K_t$ satisfying this equation.

Combining the above equations, including the social resource constraint and the binding capital requirement, it is apparent that the allocations of K_t , c_t and d_t are identical to those of the decentralized equilibrium summarized above in equations (32) through (36). Case (a) corresponds to the case of pure bank finance (also termed ‘case (a)’ in the main text), while case (b) corresponds to firms relying on both equity and bank finance (again, same label in the main text). Hence, *this constrained social planner’s problem replicates the decentralized equilibrium* when $\sigma = 0$ in the decentralized equilibrium and financial intermediation is costly.

3. Proof of Proposition 2. By assumption (10) holds: $S(T) \leq \gamma R_t^E$ for all $t \geq 0$, so $\sigma_t = 0$ in the decentralized equilibrium. The marginal effect on welfare of raising γ , without altering T , is now:

$$\frac{\partial V_0^{ci}(\theta)}{\partial \gamma} = \frac{\partial \ell_0^{ci}(\theta)}{\partial \gamma} = - \sum_{t=0}^{\infty} \beta^t \chi_t^{sp} L_t$$

where it is recalled that χ_t^{sp} is the Kuhn Tucker multiplier on the capital requirement of the social planner's problem. Using the first order conditions (d) and (c) to the social planner's problem above,

$$\chi_t^{sp} L_t = (u_d(c_t, d_t) - u_c(c_t, d_t) g_D(d_t, L_t)) L_t$$

Since the allocations of K_t , c_t and d_t are identical to those of the decentralized equilibrium, I can use the decentralized equilibrium values for the variables on the right hand side of this equation. Moreover, in the decentralized equilibrium, we have, using (2),

$$(u_d(c_t, d_t) - u_c(c_t, d_t) g_D(d_t, L_t)) L_t = u_c(c_t, d_t) (R_t^E - R_t^D - g_D(d_t, L_t)) (1 - \gamma)^{-1} d_t$$

where I have also used the result that the capital requirement binds, so that $d_t = (1 - \gamma)L_t$.

Next, I use the assumption that the economy is in steady state in period 0. Then the first order approximation of the welfare effect of an increase in γ by $\Delta\gamma$ simplifies as follows:

$$\frac{\partial V_0^{ci}(\theta)}{\partial \gamma} \Delta\gamma = -\frac{\chi_0^{sp} L_0}{1 - \beta} \Delta\gamma = -\frac{u_c(c_0, d_0) (R_0^E - R_0^D - g_D(d_0, L_0)) d_0}{(1 - \beta)(1 - \gamma)} \Delta\gamma$$

Again, it is useful to compare this to the welfare effect of a permanent change in consumption by a factor $(1 + \nu)$, given to a first order approximation in equation (25) if period 0 is a steady state. Equating the two effects and solving for ν immediately yields Proposition 2.

APPENDIX D. GENERAL DISTRIBUTION OF ε .

This appendix shows that the results in the main text do not depend on the specific distribution of ε , the shock to the return on risky loans, assumed in (3). Assumption (3) is generalized to the following:

Assumption: ε has a cumulative distribution function F_ε with bounded support $[\underline{\varepsilon}, \bar{\varepsilon}]$, with $-\infty < \underline{\varepsilon} < 0 < \bar{\varepsilon} < \infty$ (Formally, $\underline{\varepsilon} = \sup\{x \in \mathbb{R} \mid F_\varepsilon(x) = 0\}$ and $\bar{\varepsilon} = \min\{x \in \mathbb{R} \mid F_\varepsilon(x) = 1\}$). The mean of ε is equal to $-\xi$ ($\xi \geq 0$).

The assumption that the mean of ε is negative is maintained because the shock is meant to correspond to excessive risk taking. Note that F_ε need not be continuous, so (3) is a special case.

Net cash-flow to shareholders is $\left((R^L + \sigma\varepsilon)L - R^D D\right)^+$. Let $\hat{\varepsilon}$ be defined by $(R^L + \sigma\hat{\varepsilon})L - R^D D \equiv 0$. That is, $\hat{\varepsilon} = -(1/\sigma)(R^L - R^D D/L)$.

$$\begin{aligned}\mathbb{E}_\varepsilon \left[\left((R^L + \sigma\varepsilon)L - R^D D \right)^+ \right] &= \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left((R^L + \sigma\varepsilon)L - R^D D \right) F_\varepsilon(d\varepsilon) \\ &= \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left((R^L + \sigma\varepsilon)L - R^D D \right) F_\varepsilon(d\varepsilon) - \int_{\underline{\varepsilon}}^{\hat{\varepsilon}} \left((R^L + \sigma\varepsilon)L - R^D D \right) F_\varepsilon(d\varepsilon) \\ &= (R^L - \sigma\xi)L - R^D D + \sigma L \int_{\underline{\varepsilon}}^{\hat{\varepsilon}} (\hat{\varepsilon} - \varepsilon) F_\varepsilon(d\varepsilon)\end{aligned}$$

where the last equality uses the definition of $\hat{\varepsilon}$. If $\hat{\varepsilon} \leq \underline{\varepsilon}$, the integral in the last line equals zero and the expected net cash flow decreases linearly in σ . If $\hat{\varepsilon} > \underline{\varepsilon}$, the last term is strictly positive and increasing in σ , since $\hat{\varepsilon}$ is increasing in σ . It is straightforward to show that it is also convex in σ (a proof is included at the end of this appendix). Hence, the expected net cash flow to shareholders is convex in σ . Therefore, there are two candidates for the optimal choice for riskiness: 0 and $\bar{\sigma}$. By evaluating the two cases it is easy to verify that $\sigma = 0$ if and only if

$$\int_{\underline{\varepsilon}}^{\hat{\varepsilon}} (\hat{\varepsilon} - \varepsilon) F_\varepsilon(d\varepsilon) \leq \xi, \text{ with } \hat{\varepsilon} = -(1/\bar{\sigma})(R^L - R^D D/L).$$

Let ϕ_ε be defined by $\int_{\underline{\varepsilon}}^{\phi_\varepsilon} (\phi_\varepsilon - \varepsilon) F_\varepsilon(d\varepsilon) \equiv \xi$. Note that ϕ_ε exists and is unique. The definition of ξ and the assumptions that $\xi > 0$ and $\bar{\varepsilon} > 0$ imply that $\underline{\varepsilon} < \phi_\varepsilon < 0$. Restating the above condition, we have that $\sigma = 0$ if and only if

$$-(1/\bar{\sigma})(R^L - R^D D/L) \leq \phi_\varepsilon$$

With an appropriate choice of σ_{RF} (and thus $\bar{\sigma}$), ϕ_ε can be normalized to $\phi_\varepsilon = -1$. (Formally, for $\lambda \in \mathbb{R}$, let $F_{\lambda\varepsilon}$ be the c.d.f. of $\lambda\varepsilon$: $F_{\lambda\varepsilon}(x) \equiv F_\varepsilon(x/\lambda)$, for all $x \in \mathbb{R}$. Then, for any $\lambda > 0$, $(\sigma_{RF}/\lambda, F_{\lambda\varepsilon})$ presents exactly the same risky technology as $(\sigma_{RF}, F_\varepsilon)$ and $(\bar{\sigma}/\lambda, F_{\lambda\varepsilon})$ presents exactly the same opportunities to the bank as $(\bar{\sigma}, F_\varepsilon)$.⁵¹ Rescaling ε by λ in this way results in $\mathbb{E}[\lambda\varepsilon] = -\lambda\xi$ and $\phi_{\lambda\varepsilon} = \lambda\phi_\varepsilon$. The latter is true because

$$\int_{\lambda\underline{\varepsilon}}^{\lambda\phi_\varepsilon} (\lambda\phi_\varepsilon - \eta) F_{\lambda\varepsilon}(d\eta) = \int_{\lambda\underline{\varepsilon}}^{\lambda\phi_\varepsilon} (\lambda\phi_\varepsilon - \eta) F_\varepsilon(\lambda^{-1}d\eta) = \int_{\underline{\varepsilon}}^{\phi_\varepsilon} (\lambda\phi_\varepsilon - \lambda\varepsilon) F_\varepsilon(d\varepsilon) = \lambda\xi$$

⁵¹ Recall that σ/σ_{RF} is the fraction of loans made to a single risky firm and $\bar{\sigma}$ is the supervisory upper bound on σ . Rescaling ε by λ and σ_{RF} by $1/\lambda$ implies rescaling $\bar{\sigma}$ by $1/\lambda$.

where the second-to-last step performs a change of variable ($\varepsilon = \eta/\lambda$). Hence, by setting $\lambda = -1/\phi_\varepsilon > 0$ we get the desired normalization.) With that normalization, we have

$$\begin{aligned}\sigma &= 0 \text{ iff } \bar{\sigma} \leq R^L - R^D(D/L) \\ \sigma &= \bar{\sigma} \text{ otherwise ,}\end{aligned}$$

the same result as (5) in the main text. The rest of the analysis is the same as in the main text. Hence,

$$\bar{\sigma} \leq \gamma R^E$$

is a sufficient condition for no excessive risk taking (and is also necessary when the capital requirement is binding), as claimed.

Finally, the definition of ϕ_ε implies that, if $\xi = 0$, $\phi_\varepsilon = \underline{\varepsilon}$. The normalization $\phi_\varepsilon = -1$, then implies $\underline{\varepsilon} = -1$. By continuity, if $\xi \approx 0$, $\underline{\varepsilon} \approx -1$.

Proof that expected net cash flow to shareholders is convex in σ . It remains to be shown that

$$h(\sigma) \equiv \sigma \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma)} (\hat{\varepsilon}(\sigma) - \varepsilon) F_\varepsilon(d\varepsilon) , \text{ with } \hat{\varepsilon}(\sigma) = -r/\sigma$$

and $r \equiv R^L - R^D D/L$, is convex.

Proof: Let $\sigma_1 < \sigma_2$ and, for $\lambda \in (0,1)$, define $\sigma_\lambda \equiv \lambda\sigma_1 + (1-\lambda)\sigma_2$. Let $\hat{\varepsilon}_i = \hat{\varepsilon}(\sigma_i)$, for $i = 1, 2, \lambda$. Note that $\hat{\varepsilon}_1 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_2$.

$$\begin{aligned}h(\sigma_\lambda) &= \lambda\sigma_1 \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon) F_\varepsilon(d\varepsilon) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon) F_\varepsilon(d\varepsilon) \right\} + \\ &\quad (1-\lambda)\sigma_2 \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon) F_\varepsilon(d\varepsilon) - \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon) F_\varepsilon(d\varepsilon) \right\} \\ &= \lambda\sigma_1 \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon) F_\varepsilon(d\varepsilon) + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) F_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon) F_\varepsilon(d\varepsilon) \right\} + \\ &\quad (1-\lambda)\sigma_2 \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon) F_\varepsilon(d\varepsilon) + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) F_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\varepsilon - \hat{\varepsilon}_\lambda) F_\varepsilon(d\varepsilon) \right\} \\ &\leq \lambda h(\sigma_1) + \lambda\sigma_1 \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) F_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) F_\varepsilon(d\varepsilon) \right\} + \\ &\quad (1-\lambda)h(\sigma_2) + (1-\lambda)\sigma_2 \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) F_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \hat{\varepsilon}_\lambda) F_\varepsilon(d\varepsilon) \right\} \\ &= \lambda h(\sigma_1) + (1-\lambda)h(\sigma_2) + F_\varepsilon(\hat{\varepsilon}_\lambda) (\lambda\sigma_1(\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_2(\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)) \\ &= \lambda h(\sigma_1) + (1-\lambda)h(\sigma_2)\end{aligned}$$

where the last step follows from $\sigma_i \hat{\varepsilon}_i = -r$ for $i = 1, 2$ and λ . Hence,

$h(\sigma_\lambda) \leq \lambda h(\sigma_1) + (1-\lambda)h(\sigma_2)$, so $h(\sigma)$ is convex.

References

Allen, Franklin, and Douglas Gale, 2003. "Capital Adequacy Regulation: In Search of a Rationale" in *Economics in an Imperfect World: Essays in Honor of Joseph Stiglitz*, edited by R. Arnott, B. Greenwald, R. Kanbur and B. Nalebuff, Cambridge, MA: MIT Press 2003, 83-109.

Diamond, Douglas and Philip Dybvig, 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91, 401-19.

Diamond, Douglas W., and Raghuram G. Rajan, 2000. "A Theory of bank Capital", *Journal of Finance*, December.

Dewatripont, Mathias, and Jean Tirole, 1994. *The Prudential Regulation of Banks*, MIT Press, Cambridge, MA.

Feenstra, Robert, 1986. "Functional Equivalence between Liquidity Costs and the Utility of Money", *Journal of Monetary Economics* 17, pp. 73-85.

Giammarino, Ronald M., Tracy R. Lewis, and David Sappington, 1993. "An Incentive Approach to Banking Regulation", *Journal of Finance*, 48(4), pp. 1523-42.

Gorton, Gary, and Andrew Winton, 2000. "Liquidity Provision, the Cost of Bank Capital, and the Macroeconomy", working paper, University of Minnesota.

Hawke, John D., 2000. "Deposit Insurance Reform and the Cost of Bank Supervision", speech before the Exchequer Club, Washington, D.C.

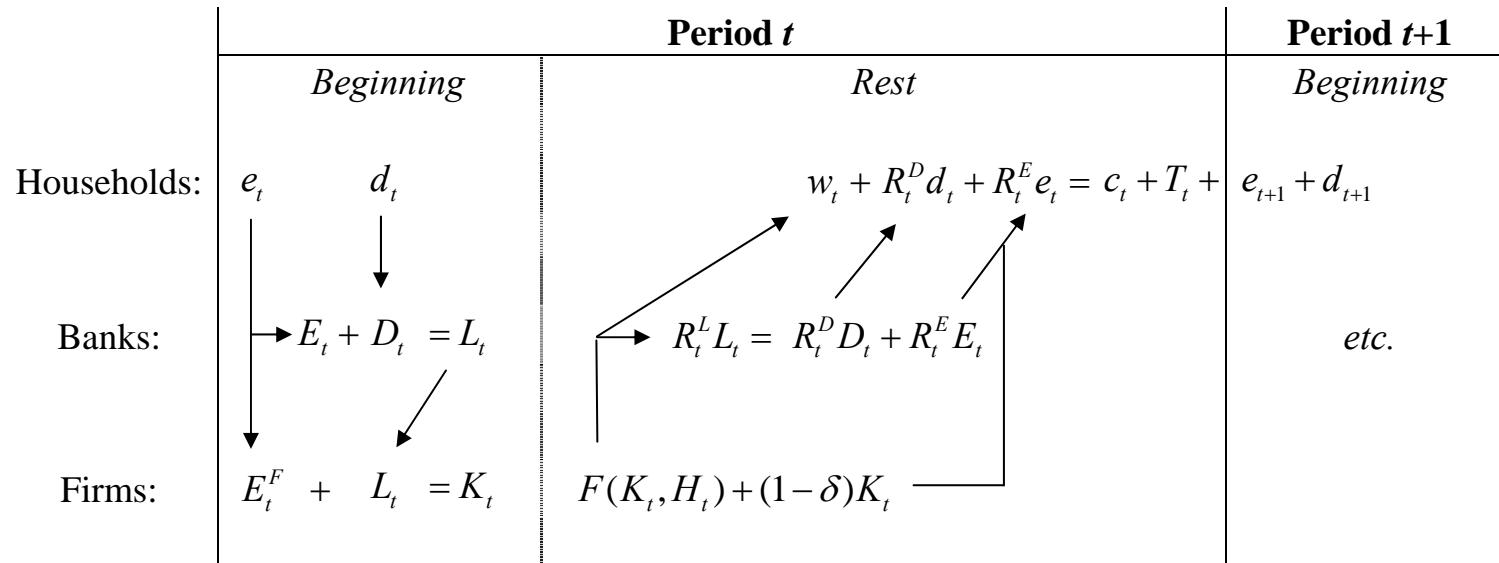
Lucas, Robert E, Jr, 2000. "Inflation and Welfare", *Econometrica*, vol. 68, no. 2, March, pp. 247-74.

Sidrauski, Miguel, 1967a. "Rational Choice and Patterns of Growth in a Monetary Economy", *American Economic Review*, May, 57(2).

Sidrausky, Miguel, 1967b. "Inflation and Economic Growth", *Journal of Political Economy*, 75 (Dec.), 798-810.

Van den Heuvel, Skander J., 2004. "The Bank Capital Channel of Monetary Policy", University of Pennsylvania, <http://finance.wharton.upenn.edu/~vdheuvel/BCC.pdf>

Figure 1. Timeline.



Arrows represent financial flows and wage payments.