Preventing Zombie Lending*

Max Bruche
Cass Business School†

Gerard Llobet
CEMFI‡

This version: October 2012

Abstract

Because of limited liability, insolvent banks have an incentive to continue lending to insolvent borrowers, in order to hide losses and gamble for resurrection, even though this is socially inefficient. We suggest a scheme that regulators could use to solve this problem. The scheme would induce banks to reveal their bad loans, which can then be dealt with. Bank participation in the scheme would be voluntary. Even though banks have private information on the quantity of bad loans on their balance sheet, the scheme avoids creating windfall gains for bank equity holders. In addition, some losses can be imposed on debt holders.

JEL codes: G21, G28, D86
keywords: Bank bail-outs, forbearance lending, recapitalizations, asset buybacks, mechanism design

*We would like to thank Juanjo Ganuza, John V. Duca, Ulrich Hege, Michael Manove, Stephen Morris, Nicola Persico, Rafael Repullo, David Ross, José Scheinkmann, Javier Suárez, Jean-Charles Rochet, and Jean Tirole, as well as our discussants, and participants at the various seminars and conferences where this paper was presented, for helpful comments. This paper originally circulated under the name "Walking Wounded or Living Dead? Making Banks Foreclose Bad Loans."

†Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK. Phone: +44 20 7040 5106. Fax: +44 20 7040 8881. Email: `max.bruche.1@city.ac.uk`.

‡CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. Phone: +34 91 429 0551. Fax: +34 91 429 1056. Email: `llobet@cemfi.es`.

1 Introduction

When too many of its borrowers turn out to be insolvent, a bank becomes insolvent. Even though continuing to lend to these insolvent borrowers tends to destroy rather than create value (because the insolvent borrowers face the wrong incentives), it also avoids crystallizing losses. An insolvent bank can therefore have incentives to continue to lend in order to hide the fact that it is insolvent, while hoping for an improvement of the situation of its insolvent borrowers. This type of gamble for resurrection is sometimes called “zombie lending,” “evergreening,” “forbearance lending,” or “extending and pretending.” If many banks engage in zombie lending, then the resulting misallocation of credit towards insolvent borrowers that should go bankrupt and are kept alive can have damaging economic consequences.

There is formal evidence that such zombie lending took place in Japan during the 1990s (Peek and Rosengren, 2005; Sekine, Kobayashi, and Saita, 2003), and that this produced substantial economic damage: Caballero, Hoshi, and Kashyap (2008) argue that keeping zombie firms alive prevented entry of more efficient ones, and caused the Japanese ‘lost decade’ of growth. For the current financial crisis, there is no formal analysis yet, but some anecdotal evidence suggests that zombie lending might be taking place. For example, in Spain, there is a concern that banks have been hiding bad loans by rolling them over.[1] In Ireland, it appears that zombie banks have kept zombie hotels alive in order to avoid crystallizing losses on loans to these hotels, which is causing major damage to the solvent competitors.[2] Also, in a recent Financial Stability Report, the Bank of England has expressed concern about potential forbearance lending in the UK, and noted that its true nature and extent is difficult to quantify due to insufficient information.[3]

When banks can hide bad loans via zombie lending, they are likely to be better informed about the true quantity of bad loans on their own balance sheet than the regulator. In a regulatory intervention, banks are likely to exploit this informational advantage in order to maximize transfers that they receive, that is, to obtain *information rents*. Avoiding such rents is important for several reasons. First, information rents are politically problematic because the public can perceive them as a reward to banks that have taken unnecessary risks. Second, they can distort ex-ante incentives of banks to screen borrowers properly. Finally, they are socially costly because of the taxation necessary to finance them.

[1] “Instead of disclosing troubled credit, many Spanish lenders have chosen to refinance loans that could still prove faulty and to report foreclosed or unsold homes as assets, often without posting their drop in market value.” See “Zombie Buildings Shadow Spain’s Economic Future,” The Wall Street Journal, September 16, 2010.

[2] See “Zombie Hotels Arise in Ireland as Recession Empties Rooms,” http://www.bloomberg.com, Aug 30, 2010.

[3] See the Bank of England’s Financial Stability Report, June 2011, especially section 2.2 and Box 2.

In this paper we suggest a scheme that regulators could use to deal with this problem. The scheme can be implemented by subsidizing the foreclosure or modification of loans, or via an asset buyback – a transaction in which the regulator buys the bad loans from banks and then forecloses or modifies them. The key insight is that, since banks have private information on quantities of bad loans, the scheme must price discriminate on quantities in order to limit the extent to which banks can exploit their private information.

This price discrimination can be implemented in various ways, for instance by allowing banks to select a two-part tariff from a menu. In the context of an asset buyback, each tariff would consist of an initial flat fee that the bank must pay to participate, and a unit price that it will then receive for each loan that it sells. Naturally, one will want to structure the menu such that higher fees are associated with higher prices. When faced with this menu, banks with a higher proportion of bad loans will select tariffs with a higher price and a higher fee. This is because they have more bad loans to sell, and therefore care more about obtaining a higher price for their bad loans. A careful structuring of fees and prices allows the reduction of information rents to bank equity holders.

In fact, we show how and when our scheme can be structured so as to afford *no* information rents to bank equity holders. In other words, the scheme makes banks solvent and prevents them from engaging in zombie lending, but banks do not benefit from participating in the scheme. Importantly, we show how the fundamental features of the problem that can cause the zombie lending in the first place, namely limited liability of banks and the risk inherent in hanging on to bad loans, are closely related to the features of the problem that allow the elimination of rents to equity holders. In addition, we discuss when losses can be imposed on debt holders, to further reduce the cost of the scheme.

In our model, banks have good and bad loans on their balance sheet. Good loans always generate a higher expected return than bad loans. Bank managers act to maximize the value of bank equity. When faced with a bad loan, bank managers must decide whether to realize a loss on the loan immediately or to delay the realization of this loss. The bank knows the size of the loss if it is realized immediately, but is uncertain about the size of the loss if it is delayed. We assume that, in expected net present value terms, delaying the realization of losses increases the likely size of the loss. This assumption can be motivated by the observation that insolvent borrowers that are given extra time because action on their loan is delayed are likely to have incentives to extract value, whatever the cost to the bank. Overall, their actions are likely to destroy value.

This choice between acting now or delaying can be interpreted as, for example, the choice between foreclosing a bad loan, or rolling it over (and foreclosing later). If a bank forecloses immediately, and seizes a building as collateral (say), it might have an idea about the price it would obtain for the building if sold now, but not if it rolled over the loan and sold it later — in which case the future recovery would depend on the evolution of property prices. Alternatively, it can be interpreted as the choice between modifying the terms of the loan to ensure that the borrower can repay, or not modifying the loan. For instance, if the bank cuts face value by 50%, it might know that the borrower will then have no problems in repaying, and hence knows exactly what loss it is incurring. If it does not modify the loan, the amount that it will ultimately recover will depend on what assets it might or might not be able to recover from the borrower.

In this context, without intervention, banks that have few bad loans foreclose or modify all of their bad loans, and banks that have many bad loans foreclose or modify none of them, and engage in zombie lending as a gamble for resurrection. This is a generic limited liability distortion.

We assume that banks have an informational advantage, in that they know the proportion of bad loans on their balance sheet, and hence know how solvent they are, but that the regulator does not.[4] In this context, it is clear that simple schemes might produce large information rents for bank equity holders. We show that the reason that the type of price-discriminating scheme described above cannot just reduce rents, but completely eliminate them is intimately related to the convexities introduced by limited liability. Global rent-elimination in an optimal contract is not something that is obtained in typical mechanism design models precisely because these tend to posit concave objective functions; in our case, the convexity introduced by limited liability makes banks want to either realize all losses immediately, or none, meaning that it is in a sense easier to induce them to realize all losses immediately. Furthermore, the convexity introduced by limited liability also affects the outside option of banks in a crucial way which makes eliminating rents incentive compatible. This second effect of limited liability can be interpreted in terms of *countervailing incentives* as discussed in the mechanism design literature (Lewis and Sappington, 1989; Maggi and Rodríguez-Clare, 1995; Jullien, 2000).

One concern is that if banks anticipate that a resolution scheme will be implemented, then this might give them weaker incentives to screen their borrowers properly going forward. This is less of a concern with our proposed scheme, precisely because we eliminate information rents. For any arbitrary proportion of bad loans, the value of equity under our scheme is exactly equal to the value of equity in the absence of intervention. This means that under our scheme, banks have incentives to be as careful in screening borrowers as in the absence of intervention. This is in contrast with alternative schemes. Consider, for example,

[4]The model could also be interpreted as describing a situation in which the regulator has received a signal on the solvency of a single bank.

a naive implementation of an asset buyback, with a single price and no participation fee. Because more insolvent banks attach a "gambling value" to their bad loans over and above fundamental value (derived from the limited liability put), the regulator necessarily has to set the price above fundamental value. This means that in the naive asset buyback, larger information rents are paid to banks with a larger proportion of bad loans. If those rents are anticipated, banks will have less incentives to screen borrowers carefully ex-ante.

In the baseline model, it can be optimal not to bail out the most insolvent banks. The reason is that in a bailout, banks need to receive transfers that compensate them for giving up the gambling value that they extract from their limited liability put, and this value is increasing and convex in the proportion of bad loans. Although the most insolvent banks have more bad loans, and therefore preventing them from zombie lending preserves more value, it can therefore also be much more costly to bail them out, and the costs can outweigh the benefits. Adding other elements that affect the cost-benefit calculation, such as a social cost of bank failure, deposit insurance, or crowding out effects (as in Caballero, Hoshi, and Kashyap, 2008), typically entails a change in the set of banks that a regulator would optimally bail out, but does not affect the fact that rents can still be eliminated.

In our scheme, debt that is initially risky becomes risk-free for the participating banks. This implies that debt holders benefit from the scheme, even when equity holders do not. This implicit rent to debt holders increases the cost of the scheme. With a slightly modified version of the scheme, we illustrate that the extent that debt holders can be made to accept losses is likely to depend on the ability of the regulator to commit to punishing debt holders who do not accept those losses, by not bailing out their banks. If the regulator can perfectly commit, the cost of the scheme can actually become negative because the regulator can appropriate the increase in value generated by stopping banks from gambling. If the regulator cannot commit at all, as it is likely to be the case in practice, the losses that can be imposed on debt holders are limited. Interestingly, the inability to fund large bailouts can create a form of commitment and help in extracting concessions from debt holders.

Although knowledge of the quantity of bad loans on any bank's balance sheet is not required to implement the optimal scheme, it does require knowledge of three key pieces of information for each bank: its leverage, the recoveries that the bank can obtain by acting immediately on bad loans, and the hypothetical distribution of future recoveries that the bank can obtain by delaying action. First, we would argue that regulators have relatively good information about bank leverage. Second, even though regulators might not know the recoveries that banks can obtain by acting on their bad loans immediately, there are implementations of the optimal scheme that generate this information. For example, in an asset buyback in which the regulator first buys bad loans, and then forecloses or modifies them, the regulator observes the recoveries on the loans, and can condition payments to banks on this additional piece of information if necessary. Third, the regulator would need to perform some calculations to forecast hypothetical future recoveries. Methodologically, these would not be very different from some of the calculations that are carried out for the "stress tests" commonly used by regulators, in which losses in different macroeconomic scenarios are forecast.[5] Although the calculations are not trivial, we believe that they do not diverge very much from the kind of calculations that bank regulators perform on a regular basis, and we therefore believe that it should be feasible to implement a version of the scheme that we describe in practice.

Finally, we consider to what extent our argument could provide insights in a case in which loans are not directly held on banks' balance sheets, but are on the balance sheets of Special Purpose Vehicles (SPVs) in securitization deals. In this case, foreclosure or modification decisions are made by the so-called servicers associated with the securitization deals. A regulator might worry both about the incentives of banks that have large positions in "toxic" securities issued by SPVs, and also directly about the incentives of servicers. At the level of servicers, we argue that for a specific type of deal (those of commercial mortgage backed securitization deals in which servicers have exposure to first loss pieces), the incentives of servicers are very similar to those of banks in our model, and our scheme could be applied almost one-for-one to servicers instead of banks. At the level of banks, regulators might use a version of our scheme to buy all tranches of toxic securities and sell them to outside investors. We also argue that a version of our scheme may be used to remove toxic securities from balance sheets, not necessarily to prevent zombie lending but to stimulate lending by eliminating debt overhang.

Related literature There is a growing literature of papers that are are motivated by the recent crisis and apply ideas from mechanism design to the problem of bailing out insolvent banks. For example, Philippon and Schnabl (forthcoming) consider a debt overhang problem. In their setting, banks differ and have private information across two dimensions: the probability of a high-payoff state of their in-place assets, and the value of their new investment opportunities. They emphasize heterogeneity along the second dimension. In the optimal intervention, banks sell warrants because the willingness to part with warrants can reveal information about the value of new investment opportunities. In contrast, we emphasize heterogeneity in the quantity of bad loans. In our optimal intervention, the

[5]Note that again, if the regulator implements the scheme in a way that allows observing the actual recoveries when loans are foreclosed or modified immediately, this information could be used to make inferences about the quality of the underlying collateral, which might be useful for estimating the recoveries that could have been obtained if action on the bad loans had been delayed.

willingness of banks to part with a given quantity of loans can reveal information about the quantity of bad loans on the bank's balance sheet.

In a paper contemporaneous to ours, Bhattacharya and Nyborg (2010) also consider a debt overhang problem. They generalize the setting of Philippon and Schnabl (forthcoming) by considering a situation in which banks not only differ in the probability of the high-payoff state of their in-place assets, but also in the size of the payoff in the low-payoff state, in a way such that in-place assets of different banks can be ranked in a first-order stochastic dominance sense. They then show that a menu of equity injections can separate the banks, and that under a monotonicity condition on payoffs and probabilities, information rents can be eliminated.

In the setup of Bhattacharya and Nyborg (2010), banks take no decision, which simplifies the contracting problem. In contrast, we consider a setup in which banks take a decision (whether or not to foreclose, modify, or sell bad loans). In our setting, optimal schemes make transfers to banks directly conditional on bank decisions. The type of unconditional equity injection considered by Bhattacharya and Nyborg (2010) cannot be optimal in our setting, because it only influences bank behaviour very indirectly, by affecting solvency.[6]

Also, relative to Bhattacharya and Nyborg (2010) we put more structure on the assets of a bank, and describe solvency as being related to the proportion of bad loans. Our counterpart to their monotonicity condition is the requirement that non-participation values of equity need to be convex in the proportion of bad loans. In our setup, this condition is always satisfied because of how we relate the returns on banks' assets to the proportion of bad loans on its balance sheet, and the presence of limited liability.

A third important difference of our paper with respect to theirs is that they assume that the regulator will always want to bail out all banks, whereas we show how the set of banks that is optimally bailed out can vary with the particular choice of welfare function, even though the nature of the optimal contract does not vary.

Three related papers are those of House and Masatlioglu (2010), Philippon and Skreta (2012), and Tirole (2012). They consider a situation in which the main problem is one of adverse selection in markets relevant for the funding of banks. Via some scheme, the regulator provides an alternative source of funds. The participation decisions of banks affect which banks will remain funded by the market, and consequently the degree of adverse selection in this market. Since the market for funding is the outside option of all banks, their participation constraint in the scheme becomes endogenous. The optimal scheme needs to

[6]Interestingly, Giannetti and Simonov (forthcoming) show that many of the equity injections carried out in Japan were relatively small and were not conditional on a change of behaviour of banks, and therefore in many cases did not prevent further zombie lending.

take this into account. We abstract from such problems here to focus on our core message.

There is also a literature that views asset buybacks as a solution to the problem of fire-sale discounts. Diamond and Rajan (2011) describe how a regulator can ensure bank liquidity by buying assets from banks at prices above those that current private buyers are willing to pay, but below the fundamental value of the asset. Gorton and Huang (2004) show that it can be more efficient for the government rather than the private sector to provide liquidity by buying up bank assets. In the context of providing liquidity via asset purchases, the papers of Ausubel and Cramton (2008) and Klemperer (2010) have proposed auction designs that aim to to prevent paying more than fundamental value for the assets. In contrast, in our model, asset buybacks are a solution to the problem of inefficient gambling for resurrection by banks. Since distressed banks want to gamble, anyone attempting to buy a bad asset will necessarily have to pay more than fundamental value in order for such a bank to part with the bad asset. As we show, overpaying for the bad assets does not necessarily imply windfall gains for bank equity holders.

Many papers, including those of Mitchell (1998), Corbett and Mitchell (2000), and Mitchell (2001) examine models in which the proportion of bad debt on a bank's balance sheet is private information and bank managers can hide bad loans via rolling them over. In the same type of setting, Aghion, Bolton, and Fries (1999) argue that there is a tradeoff between having "tough" closure policies for banks, which gives incentives to hide problems ex-post but provides incentives not to take risks ex-ante, and having "soft" closure policies for banks, which does not give incentives to hide problems ex-post, but provides incentives to take risks ex-ante. Although not the main focus of their paper, they also sketch a second-best scheme that involves transfers conditional on the liquidation of non-performing loans.

Our paper is also related to the general mechanism design literature. The two-part tariff implementation of our optimal contract turns out to be mathematically similar to the original problem of Baron and Myerson (1982), except that we have a type-dependent outside option. This creates what Lewis and Sappington (1989) referred to as "countervailing incentives". In our case, though, the type-dependent outside option is not concave but convex in types, due to the convexity introduced by limited liability, which has been shown to imply that rents can be eliminated for a range of types (Maggi and Rodríguez-Clare, 1995; Jullien, 2000). In addition, in our case the agents' objective function is convex in the decision, allowing us to actually eliminate rents for *all* types.

In section 2, the basic model is set up. In section 3, we present the optimal contract and various implementations. Section 4 examines to what extent different implementations of the optimal contract are robust to a situation in which banks can pretend that good loans are bad in order to obtain higher transfers. Section 5 studies under which conditions losses can be imposed on debt holders. Section 6 studies other social welfare functions. Section 7 discusses how the information requirements of the scheme could be overcome in practice. Section 8 addresses to what extent the model can provide insights in a situation in which loans are not held directly on banks' balance sheets, but on the balance sheets of SPVs in securitization deals. Section 9 concludes. All proofs are in the appendix.

2 The model

Consider an economy with two dates $t = 1, 2$. There is no discounting across periods. There exists a continuum of risk-neutral banks, that operate under limited liability and maximize the expected value of their equity. All banks have debt with face value D, due to be paid at $t = 2$, where $0 < D < 1$. All banks have a measure 1 of loans. Each loan has a face value of 1. Loans can be either good or bad. At date $t = 1$, each bank learns what proportion θ of its loans are bad loans, and what proportion $1 - \theta$ of its loans are good loans. The proportion θ varies across banks and is private information. The distribution of θ in the population of banks is denoted as $\Psi(\theta)$ with density $\psi(\theta)$.

At $t = 1$, after learning θ, banks can decide on which amount γ of bad loans they want to take immediate action, where $\gamma \in [0, \theta]$. On the remaining bad loans, an amount $\theta - \gamma$, action is delayed. Any bad loan on which immediate action is taken at $t = 1$ produces a recovery of $\rho < 1$. As mentioned in the introduction, we consider two different interpretations. In the first interpretation, taking immediate action means foreclosing after which the bank obtains a recovery of ρ at $t = 1$. We assume that in this case, the bank cannot pay dividends at $t = 1$ such that the proceeds from foreclosure are carried forward until $t = 2$. In the second interpretation, taking immediate action means modifying the loan by cutting face value from 1 to ρ. This ensures that the borrower will be able to repay this amount for sure at $t = 2$. For ease of exposition, we only refer to the first interpretation of the decision (foreclose/ roll over) for the remainder of this section.

At $t = 2$, any good loan pays off 1. For bad loans on which action was delayed at $t = 1$ the payoff is realized now, producing a random recovery of ε. The realization of ε is the same for all such loans of a given bank. The distribution of ε has full support in $[0, 1]$ and is denoted by $\Phi(\varepsilon)$, and its density by $\phi(\varepsilon)$. We assume that $E[\varepsilon] < \rho$, such that rolling over loans destroys net present value.[7]

If in the second period the realized ε is sufficiently low, a bank will not be able to repay

[7] As noted in the introduction, insolvent borrowers whose loans are rolled over (or not modified) are likely to have incentives to extract value, whatever the cost to the bank. This is likely to destroy value overall, as could be demonstrated via a model in which the borrower has to make an effort choice to maintain the value of collateral.

its existing debt. A bank that forecloses an amount of bad loans γ will survive if

$$1-\theta+(\theta-\gamma)\varepsilon+\gamma\rho>D.$$

That is, a bank will survive if it can repay D in full with the return of the good loans together with the return from bad loans that have been rolled over – which depends on the realized ε – and the return from the foreclosed loans. In other words, the bank will be able to repay D as long as the realized ε is sufficiently high, or if

$$\varepsilon \geq \bar{\varepsilon}_0 \equiv \frac{\theta-\gamma\rho-(1-D)}{\theta-\gamma}. \tag{1}$$

As expected, a lower proportion of bad loans, a lower debt level, and a higher recovery upon foreclosure will increase the probability that the bank survives.

We can now write the expected value of equity of a bank that holds bad loans θ as

$$\int_{\bar{\varepsilon}_0}^{1} \left(1-\theta+(\theta-\gamma)\varepsilon+\gamma\rho-D\right)\phi(\varepsilon)d\varepsilon. \tag{2}$$

As it turns out, the value of equity is convex in γ due to the bank's limited liability. It implies that banks are interested in either foreclosing all bad loans or none. In particular, banks with few bad loans foreclose all bad loans ($\gamma=\theta$), and banks with many bad loans foreclose no bad loans ($\gamma=0$). The intuition for this result is straightforward. Banks that are likely to survive (low θ) have a valuation of rolled-over bad loans that is close to their true expected value, and hence prefer to foreclose. Banks that are not very likely to survive (high θ) have a valuation of rolled-over bad loans that only reflects their large positive returns in the state in which they survive, and hence do not foreclose. This is the typical gambling for resurrection behavior, and we will therefore refer to the banks that roll over their bad loans (do not foreclose) as *gambling banks*. We let $\hat{\theta}$ denote the critical value of θ above which banks will gamble.

Below, we will let

$$\pi_0^G(\theta)=\int_{1-(1-D)/\theta}^{1}(1-\theta+\theta\varepsilon-D)\phi(\varepsilon)d\varepsilon \tag{3}$$

denote the value of equity when gambling ($\gamma=0$), and hence $\bar{\varepsilon}_0=1-(1-D)/\theta$, and

$$\pi_0^F(\theta)=\max(1-\theta+\theta\rho-D,0) \tag{4}$$

denote the value of equity when foreclosing ($\gamma=\theta$). In terms of $\pi_0^G(\theta)$ and $\pi_0^F(\theta)$, the value of equity, taking into account that banks will choose γ optimally, can then be written as

$$\pi_0(\theta)=\max(\pi_0^G(\theta),\pi_0^F(\theta)). \tag{5}$$

Figure 1 illustrates this discussion, and Lemma 1 summarizes it formally.

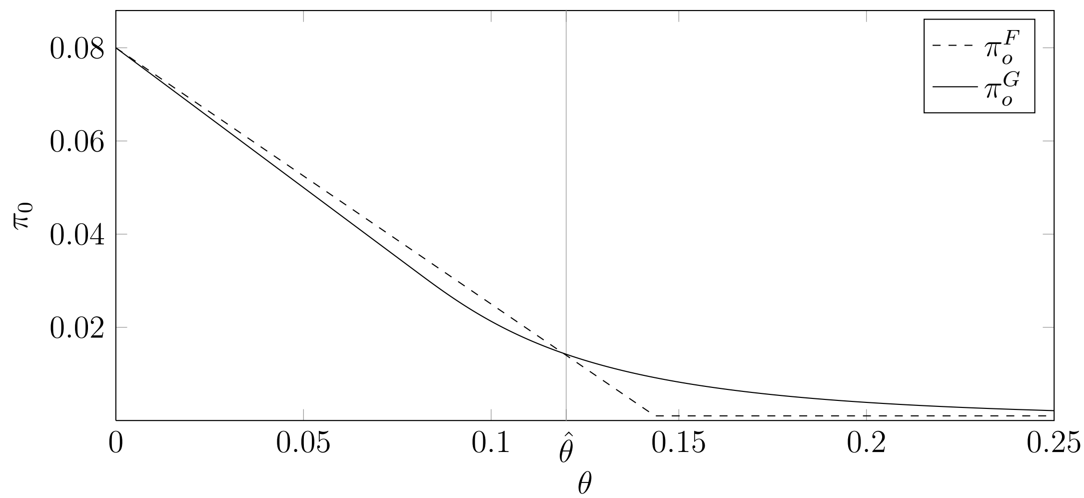

Figure 1: Equity value as a function of θ

Equity values for banks as a function of θ when banks foreclose (dashed line, $\pi_0^F(\theta)$), and when banks gamble (solid line, $\pi_0^G(\theta)$). Banks choose whichever is higher. Banks with $\theta > \hat{\theta}$ gamble, and banks with $\theta < \hat{\theta}$ foreclose. Parameters are $1-D = 0.08$, $\rho = 0.45$, and $\varepsilon \sim \text{Beta}(2,3)$, which implies $E[\varepsilon] = 0.40$.

Lemma 1. *The value of equity is convex in γ. As a consequence, a bank with a proportion of bad loans θ will decide to foreclose an amount $\gamma(\theta)$ given by*

$$\gamma(\theta) = \begin{cases} \theta & \text{if } \theta \leq \hat{\theta}, \\ 0 & \text{if } \theta > \hat{\theta}, \end{cases}$$

where $\hat{\theta}$ is defined as the (finite) value of $\theta > 0$ that solves

$$\pi_0^F(\theta) = \pi_0^G(\theta).$$

Below, we will focus on the interesting case of $\hat{\theta} < 1$ in which some banks have incentives to gamble.[8] We will also introduce a risk-neutral regulator that aims to influence the decisions of banks in order to maximize welfare.

To afford an informational advantage to banks vis-a-vis the regulator, we assume that a bank knows its θ whereas the regulator only knows the distribution of θ in the population, $\Psi(\theta)$. Furthermore, the regulator will neither observe the value of assets of a bank at $t = 2$, nor the realization of ε. This means that the regulator will not be able to indirectly infer the proportion of bad loans on a bank's balance sheet. We will assume, though, that the amount of bad loans which are foreclosed (or modified), γ, is observable and verifiable, and focus on contracts in which a bank takes action on an amount γ in exchange for a transfer T that

[8] A sufficient condition that ensures that $\hat{\theta} < 1$ would be that $\rho < D$, which ensures that banks with $\theta = 1$ obtain a value of equity of 0 from taking immediate action on all bad loans.

may or may not depend on γ. This includes, for example, contracts that pay a subsidy per foreclosed (or modified) loan, or a buyback scheme in which the regulator sets up a special purpose vehicle that buys bad loans from a bank and then forecloses (or modifies).

Note also that in our basic setup, we do not allow banks to foreclose or modify good loans. For the discussion of the case where this is possible, see Section 4.

3 The regulator's scheme

In the model described in the previous section, a bank with a large proportion of bad loans has insufficient incentives to take immediate action on its bad loans, even though delaying action destroys net present value. This destruction of net present value is socially suboptimal, and a regulator can intervene to prevent it.[9]

In this section, we first state the general optimal contracting problem that the regulator faces (Subsection 3.1). The solution to this problem (described in Subsection 3.2), involves asking participating banks to foreclose all of their bad loans, and paying a transfer that makes them just indifferent between participating or not, such that no information rents are awarded. We then present several alternative implementations of the optimal scheme in Subsection 3.3. In the context of a particular implementation via two-part tariffs, we illustrate the role of the three key properties of the model that allow rent elimination. One of these is that non-participation values of equity have to be convex in the proportion of bad loans, which can be related to *countervailing incentives* as discussed in the mechanism design literature (Lewis and Sappington, 1989; Maggi and Rodríguez-Clare, 1995).

3.1 The regulator's problem

Because the amount of loans which a bank forecloses, γ, is observable and verifiable, the regulator can transfer resources to the bank contingent on this variable, $T(\gamma)$. As usual, given the private information on θ, it is more convenient to consider direct revelation mechanisms under which a bank of type θ truthfully reports its type, and is then assigned a contract under which it forecloses an amount $\gamma(\theta)$, and in return receives a net transfer $T(\theta)$ at

[9]Could Coasian bargaining between private parties without the involvement of the regulator solve the gambling problem by a reorganization of the capital structure (Haugen and Senbet, 1978)? Here, the fact that equity holders have private information can mean that such negotiations might not take place, as in Giammarino (1989). In addition, it is reasonable to believe that banks would have incentives ex-ante to choose debt structures which would make such ex-post bargaining impossible, as argued by Bolton and Scharfstein (1996). That is, even though in the baseline model there are no externalities, it is plausible to assume that private parties would not solve the gambling problem ex-ante or ex-post.

$t = 2$.[10]

Banks facing a menu of contracts will choose the one that maximizes the value of their equity. We will denote the value of equity of a participating bank of type θ that reports type θ^R as $\Pi(\theta, \theta^R)$, given by

$$\Pi(\theta, \theta^R) = \int_{\bar{\varepsilon}}^{1} \left[1 - \theta + \left(\theta - \gamma(\theta^R)\right)\varepsilon + \gamma(\theta^R)\rho - D + T(\theta^R)\right] \phi(\varepsilon) d\varepsilon, \tag{6}$$

where

$$\bar{\varepsilon} = \frac{\theta - \gamma(\theta^R)\rho - (1 - D) - T(\theta^R)}{\theta - \gamma(\theta^R)}. \tag{7}$$

Since we consider schemes with voluntary participation, the net transfer $T(\theta)$ for a bank of type θ will have to be non-negative for that bank to participate, and might have to be positive for that bank to take immediate action on some quantity of bad loans. This implies that, in general, the scheme will not be costless. We assume that each dollar that the regulator transfers to a bank generates an associated dead-weight loss $\lambda > 0$. This loss arises, for example, if in order to finance this scheme the government needs to rely on distortionary taxation. Thus, for a given amount of foreclosed loans, the regulator will be interested in minimizing the cost of the rescue scheme.

We can then state the formal problem as follows:

$$\max_{\gamma(\theta), T(\theta)} \int_0^1 \left[1 - \theta + \theta E[\varepsilon] + (\rho - E[\varepsilon])\gamma(\theta) - \lambda T(\theta)\right] \psi(\theta) d\theta, \tag{W}$$

subject to

$$\Pi(\theta, \theta) \geq \Pi(\theta, \theta^R), \quad \forall \theta, \theta^R \tag{IC}$$

$$\Pi(\theta, \theta) \geq \pi_0(\theta), \quad \forall \theta \tag{PC}$$

$$0 \leq \gamma(\theta) \leq \theta.$$

These equations can be interpreted as follows. The objective function, (W), states that the regulator chooses the schedules $\gamma(\theta)$ and $T(\theta)$ to maximize expected welfare. The contribution of a given bank to welfare corresponds to the total value of its assets, which will be divided between its equity holders and debt holders at $t = 2$, net of the deadweight loss associated with the transfers it receives. The total value of the bank's assets are maximized when it forecloses. The main trade-off here is therefore between inducing foreclosure in order to maximize the value of assets, versus the deadweight loss associated with the transfers that induce foreclosure.

[10] We restrict ourselves to deterministic mechanisms. From a purely technical point of view, stochastic mechanisms that improve welfare exist, but they are very implausible.

The menu of contracts that the regulator offers has to induce banks to truthfully report their type, producing the incentive compatibility constraint, (IC). It also has to lead to at least the same value of equity as when not participating, producing the participation constraint (PC).

3.2 The optimal contract

The optimal contract involves paying positive transfers to a set of banks that would not foreclose in the absence of the scheme, to induce them to foreclose. We will here say that a bank *participates* in the scheme if it chooses to foreclose only in order to obtain this positive transfer, and would not foreclose in the absence of the scheme. Banks that do not receive a positive transfer do not *participate* and take their privately optimal action. In particular, for some banks with a low proportion of bad loans ($\theta < \hat{\theta}$), this will mean foreclosing anyway.

We will show that the optimal contract involves the elimination of all information rents of participating banks. This implies that banks will obtain a value of equity from participating and foreclosing that is exactly equal to the value of equity that they would obtain from staying outside the scheme. Since for participating banks, taking the privately non-optimal action (foreclosing) decreases equity value, the transfer has to just compensate for this *loss from foreclosing*, which we define as

$$\Delta\pi_0(\theta) := \pi_0(\theta) - (1 - \theta + \theta\rho - D). \tag{8}$$

Notice that in the expression for $\Delta\pi_0(\theta)$, the part $1 - \theta + \theta\rho - D$ may be negative. For a bank that has a value of total assets when foreclosing $1 - \theta + \theta\rho$ less than the face value of debt D, any transfer made to the bank needs to be used to satisfy the claim of debt holders first, before any remainder can go to equity holders. Of course, unless this remainder is positive, bank managers that act in the interest of equity holders will not in general want to participate in the scheme.

For $\theta < \hat{\theta}$, $\pi_0(\theta) = \pi_0^F(\theta)$, which, by (4), obviously implies that $\Delta\pi_0(\theta) = 0$. For banks that were already foreclosing absent the scheme, there is no loss from foreclosing. For $\theta > \hat{\theta}$, since $\pi_0(\theta) = \pi_0^G(\theta) > 1 - \theta + \theta\rho - D$, and $\pi^G(\theta)$ is decreasing, convex and has a slope bigger than -1, it follows that $\Delta\pi_0(\theta)$ is positive, increasing, and convex. For banks that were gambling absent the scheme, the loss from foreclosing is positive, increasing, and convex in the proportion of bad loans on their balance sheet.

We can now state that a scheme under which banks foreclose all bad loans while receiving a transfer exactly equal to the loss from foreclosing satisfies all the constraints in the regulator's problem:

Proposition 1. *The contract* $\{\gamma(\theta) = \theta, T(\theta) = \Delta\pi_0(\theta)\}$ *satisfies the participation constraint* (PC) *and the incentive compatibility constraint* (IC)*.*

The contract described in the lemma is one under which all banks foreclose, banks with a low proportion of bad loans ($\theta < \hat{\theta}$) that would foreclose outside the scheme receive no transfer (since for them, $\Delta\pi_0(\theta) = 0$), and only banks with a high proportion of bad loans ($\theta > \hat{\theta}$) that would gamble outside the scheme receive a positive transfer. The transfer that banks receive just offsets the loss from foreclosing, and all banks are therefore exactly as well off under the scheme as when not participating, that is to say they do not receive any information rents.

It is fairly obvious that the contract satisfies the participation constraint (PC), but less obvious that it satisfies the incentive compatibility constraint (IC). The proof for this hinges on the fact that limited liability introduces convexities — it makes the value of equity convex in the quantity of foreclosed loans, and makes the non-participation value of equity convex in the proportion of bad loans. At the same time, the fact that limited liability makes the value of equity convex in the quantity of foreclosed loans implies that the standard (first-order) approach used to characterize the set of incentive compatible contracts cannot be applied here. This means that it is hard to illustrate incentive compatibility in the context of the general optimal contract without delving into technical detail. Our strategy will therefore be to postpone a discussion of the intuition for incentive compatibility until the next subsection, where we can discuss it more easily in the context of the two-part tariff implementation of the optimal contract.

The fact that the contract in Proposition 1 is incentive compatible implies that an arbitrary set of banks can be selected to participate.

Corollary 1. *Let* $\Theta_P \subseteq [\hat{\theta}, 1]$ *denote an arbitrary set of* participating *banks. Then consider the contract*

$$\gamma^*(\theta) = \begin{cases} \theta & \text{for } \theta \in \Theta_F \\ 0 & \text{for } \theta \notin \Theta_F \end{cases}, \quad T^*(\theta) = \begin{cases} \Delta\pi_0(\theta) & \text{for } \theta \in \Theta_P \\ 0 & \text{for } \theta \notin \Theta_P \end{cases}, \tag{9}$$

where $\Theta_F = \{\theta : (\theta < \hat{\theta}) \cup (\theta \in \Theta_P)\}$ *denotes the set of banks that foreclose. Under this contract, the incentive compatibility constraint* (IC) *is satisfied for all banks, and the participation constraint* (PC) *is satisfied for all banks with equality.*

Again, it is fairly obvious that the contract in Corollary 1 satisfies the participation constraint (PC) with equality. To see that it is incentive compatible, start from a situation in which all banks that were not foreclosing absent the scheme participate, $\Theta_P = [\hat{\theta}, 1]$. This then replicates the contract in Proposition 1. We know that this is incentive compatible, and leaves all banks at their participation constraint. If we now simply eliminate contracts corresponding to some $\theta \in \Theta_P$, the banks whose point on the contract has been deleted will now prefer not to participate: By incentive compatibility of the full contract and the fact that the full contract satisfied the participation constraint with equality, they cannot obtain a higher value than their non-participation value by picking a point on the reduced contract not intended for them. Hence the reduced contract is still incentive compatible, and again leaves all banks at their participation constraints. The importance of this result is that a regulator can choose an arbitrary set of banks that should participate, and still eliminate rents.

We can now state the optimal contract:

Proposition 2. *The optimal contract* $\{\gamma^*(\theta), T^*(\theta)\}$ *that solves* (W) *subject to* (IC) *and* (PC) *is given by*

$$\gamma^*(\theta) = \begin{cases} \theta & \textit{for } \theta \in \Theta_F \\ 0 & \textit{for } \theta \notin \Theta_F \end{cases}, \quad T^*(\theta) = \begin{cases} \Delta\pi_0(\theta) & \textit{for } \theta \in \Theta_P \\ 0 & \textit{for } \theta \notin \Theta_P \end{cases}, \tag{10}$$

where $\Theta_P = [\hat{\theta}, \min(\theta^*, 1)]$ *denotes the set of banks that optimally participate and* $\Theta_F = \{\theta : (\theta < \hat{\theta}) \cup (\theta \in \Theta_P)\}$ *denotes the set of banks that foreclose. Here,* θ^* *solves*

$$(\rho - E[\varepsilon])\theta^* \equiv \lambda\Delta\pi_0(\theta^*) \textit{ and } \theta^* \geq \hat{\theta}. \tag{11}$$

This contract takes the form of the contract described in Corollary 1. Again, information rents are eliminated. The optimal set of participating banks is chosen by comparing the benefit of preventing a bank with a proportion of bad loans θ from gambling, which is the resulting increase in net present value $(\rho - E[\varepsilon])\theta$, and the cost, which is the deadweight loss associated with the required transfer, $\lambda\Delta\pi_0(\theta)$. As expected, an increase in the cost of public funds λ results in a smaller set of banks that participate.

Figure 2 illustrates why the optimal contract prescribes that banks with a higher proportion of bad loans might not be made to participate: While the benefit of having a bank participate is increasing and linear in the proportion of bad loans, the cost is increasing and convex. This reflects the fact that as one considers more and more insolvent banks, these require increasingly higher transfers in order to participate and give up their "limited liability put" (as reflected in the convexity of the loss from foreclosing). This makes the participation of very insolvent banks very expensive.

The result regarding which banks optimally participate may change with other specifications of the welfare function. If, for example, bank failures generate a significant externality, the optimal contract could prescribe that banks with very large proportions of bad loans

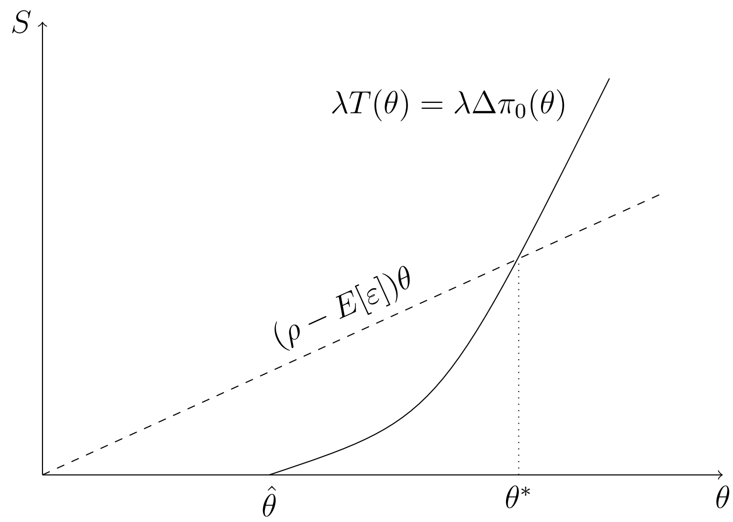

Figure 2: The optimal contract
Social benefit and losses from foreclosing for different values of θ. Banks with a proportion of bad loans between $\hat{\theta}$ and θ^* decide to participate. Banks foreclose all their bad loans if and only if $\theta < \theta^*$.

θ, which are more likely to fail, should also participate. We discuss this and other cases in Section 6. In general, in these situations, information rents can still be eliminated, as indicated by Corollary 1.

Finally, it is also useful to discuss the implications that the optimal contract has for the incentives of banks to carefully screen borrowers ex-ante. For the sake of the argument, suppose that more effort in screening borrowers ex-ante leads to an ex-post draw from a better distribution of θ in the first order stochastic dominance sense. We can intuitively see that the higher the value of equity that banks obtain for low values of θ and the lower the value of equity for high values of θ, the stronger are the incentives to exert effort. Notice that compared to the case without intervention, our mechanism provides identical incentives, since for any arbitrary value of θ, the value of equity is the same in both cases. This result is in contrast with what occurs with standard asset buybacks: If there is a single fixed price per bad loan sold (and no participation fee), information rents are granted to banks with higher ex-post values of θ. If banks anticipate this, they will respond by reducing their effort to screen borrowers ex-ante.[11]

[11] In the class of schemes with voluntary participation, the only general way of improving on the incentives produced by our scheme would be to reward banks that end up having a low proportion of bad loans with positive information rents. It can be shown that due to global incentive compatibility constraints, this necessarily also implies paying positive (although smaller) information rents to all banks that have a larger proportion of bad loans, reducing the appeal of such a mechanism. Under schemes with mandatory participation, incentives could be improved without necessarily increasing cost, but the improvement would be limited by the non-concavity of participation profits.

3.3 Implementing the optimal contract

We now discuss several alternative implementations of the optimal contract. We start with a two-part tariff implementation of a foreclosure/ modification subsidy, which we discuss in some detail by solving the corresponding contracting problem. The two-part tariff implementation allows interpreting the role of various properties of the problem that make our rent-eliminating optimal contract incentive compatible. We then also briefly discuss a non-linear foreclosure/ modification subsidy, and a two-part tariff asset buyback implementation.

The optimal contract in Proposition 2 can be implemented via a two-part tariff foreclosure/ modification subsidy: Suppose the regulator offers a menu of two-part tariffs, where each two-part tariff consists of (i) a (positive) subsidy s that the bank receives per loan that it forecloses or modifies, and (ii) a (positive) participation fee F that the bank promises to pay. Banks do not have to commit to foreclosing or modifying a specific amount, and can privately choose the amount of loans they want to foreclose or modify. In this scheme, the role of the subsidy will be to induce banks to foreclose or modify, and the role of the fee will be to claw back (some or all of) the increase in the value of equity of a bank that is derived from the subsidy.

As before, it is more convenient to consider direct revelation mechanisms under which a bank with type θ is meant to truthfully report its type and then receive the contract $(s(\theta), F(\theta))$. According to this notation, a bank that reports a type θ^R accepts to pay a fixed fee $F(\theta^R)$ in return for a subsidy $s(\theta^R)$ per foreclosed or modified loan and, thus, receives a net transfer $T(\gamma) = s(\theta^R)\gamma - F(\theta^R)$, that indirectly depends on the amount γ that the bank chooses to foreclose or modify under the tariff.

Consider a bank with a proportion of bad loans $\theta > \hat{\theta}$, that decides to participate in the scheme and picks the contract indexed by θ^R, and that subsequently forecloses or modifies a proportion γ of bad loans. In that case, the counterpart of the expected value of equity (6) under this scheme is

$$\Pi(\theta, \theta^R) = \max_{\gamma} \int_{\bar{\varepsilon}(\gamma)}^{1} \left[1 - \theta + (\theta - \gamma)\varepsilon + \gamma(\rho + s(\theta^R)) - D - F(\theta^R)\right] \phi(\varepsilon) d\varepsilon, \tag{12}$$

where

$$\bar{\varepsilon}(\gamma) = \frac{\theta - \gamma(\rho + s(\theta^R)) - (1 - D - F(\theta^R))}{\theta - \gamma}.$$

As before, it is easy to see that, due to limited liability, the value of equity is convex in γ, leading to a corner solution. Under the scheme a bank will either foreclose or modify as many bad loans as it can ($\gamma = \theta$), or not foreclose any ($\gamma = 0$). In addition, notice that a bank will never want to participate just to pay a positive fee, and not receive any subsidy in return. Thus, a participating bank will want to foreclose or modify *all* bad loans. Also, since it can always get a positive equity value by not participating, the value of equity from participating must always be strictly positive. This observation is summarized in the next remark.

Remark 1. *Under a menu of two part-tariffs with positive fees, a participating bank with a proportion of bad loans θ will find it optimal to foreclose all of its bad loans, that is, to choose $\gamma = \theta$.*

This allows us to considerably simplify the expression for the value of equity from participating. For a bank of type θ that picks the contract indexed by θ^R, this is

$$\Pi(\theta, \theta^R) = 1 - \theta + (\rho + s(\theta^R))\theta - D - F(\theta^R). \tag{13}$$

The participation constraint (PC) and the incentive compatibility constraint (IC) for the two-part tariff case can now be stated in terms of this expression.

In the rest of our discussion, it will be convenient to denote the information rents as $U(\theta)$, understood as the increase in the value of equity that a bank obtains when it participates and chooses the contract intended for its type, over the value of equity when it does not participate. That is,

$$U(\theta) \equiv \Pi(\theta, \theta) - \pi_0(\theta). \tag{14}$$

Obviously, for a bank with type θ to participate, $U(\theta) \geq 0$.

Inserting the expression for $\Pi(\theta, \theta)$ we can also express the information rents as

$$U(\theta) = \underbrace{s(\theta)\theta - F(\theta)}_{T(\theta)} - \underbrace{(\pi_0(\theta) - (1 - \theta + \theta\rho - D))}_{\Delta\pi_0(\theta)}. \tag{15}$$

In words, this states that the information rents of a bank with type θ will consist of the net transfer it receives, minus the *loss from foreclosing*, as defined in (8).

This two-part tariff problem is well-behaved, meaning that we can apply the standard first-order approach to identify conditions that characterize incentive compatible contracts. We state these in terms of the information rents.

Lemma 2. *Necessary and sufficient conditions for a two-part tariff scheme $\{s(\theta), F(\theta)\}$ to be incentive compatible are*

i) monotonicity: $s(\theta)$ is non-decreasing,
ii) local optimality:

$$\frac{dU(\theta)}{d\theta} = s(\theta) - \frac{d\Delta\pi_0(\theta)}{d\theta}. \tag{16}$$

The proof for these conditions is standard and hence omitted.[12] The first part of Lemma 2 can be interpreted as stating that banks with more bad loans should receive higher subsidies under an incentive compatible scheme. Of course, the higher subsidies will have to be associated with higher fees. Intuitively, banks with more bad loans care more about the size of the subsidy and will choose to pay a high fee and receive a high subsidy, whereas banks with a low proportion of bad loans will then choose to pay a low fee and receive a low subsidy. The second part of Lemma 2 can be interpreted as stating that to induce truth-telling, the regulator has to provide information rents that vary with the proportion of bad loans θ. The two components of the expression reflect two *countervailing incentives* that banks face, to both overstate as well as understate their type, which change with θ, as we now describe.

First, suppose the loss from foreclosing $\Delta\pi_0(\theta)$ were constant, such that the second term in (16) would be zero for all θ. Then, since the subsidy $s(\theta)$ must be positive, information rents $U(\theta)$ would have to be higher for banks with higher θ. This is because banks with high θ would otherwise understate their type, to pretend that they benefit less from the positive subsidy and in this way manage to pay a lower fee to the regulator. This incentive to understate is stronger the larger is $s(\theta)$.

Second, suppose the subsidy $s(\theta)$ were zero for all θ. Then, since the loss from foreclosing $\Delta\pi_0(\theta)$, is increasing in θ (for $\theta > \hat{\theta}$), information rents $U(\theta)$ would have to be higher for banks with lower θ. This is because banks with low θ would otherwise overstate their type, to pretend that they are incurring larger losses from foreclosing (or modifying) and in this way manage to pay a lower fee to the regulator. This incentive to overstate is larger the larger is $\frac{d\Delta\pi_0(\theta)}{d\theta}$.

The incentives to overstate and understate are in conflict, of course. A regulator that is interested in minimizing the cost of the scheme can pick $s(\theta)$ to play off the incentives for banks to overstate against the incentives to understate, in order to reduce information rents, subject to the constraints that $s(\theta)$ needs to be increasing, and $U(\theta)$ cannot be negative.

Since $\Delta\pi_0(\theta)$ is a weakly convex function of θ, the incentives to overstate are non-decreasing in θ. Intuitively, a very insolvent bank could try to extract more from the regulator by claiming to have an extra bad loan than a slightly insolvent bank could, because for the very insolvent bank, the value of the "limited liability put" is much more sensitive to its solvency. Fortunately, we can match these incentives to overstate that are non-decreasing with incentives to understate that are non-decreasing, by picking a non-decreasing function $s(\theta)$ so that the incentives to understate and overstate exactly cancel out, and leave information

[12] As highlighted in Remark 1, for positive transfers $F(\theta)$, participating banks foreclose all bad loans and hence participation profits are determined by (13). Starting from this expression, the conditions can be derived in the standard way (see for example the description in Fudenberg and Tirole (1991), section 7.3).

rents constant. In order to minimize information rents, the regulator can then choose fees that set the constant level as $U(\theta) = 0$.[13]

Following this discussion, we know that if the regulator offers the following menu of two part tariffs,

$$s^*(\theta) = \frac{d\Delta\pi_0(\theta)}{d\theta}, \tag{17}$$

$$F^*(\theta) = -\Delta\pi_0(\theta) + \theta s^*(\theta), \tag{18}$$

both conditions in Lemma 2 are satisfied: The subsidy $s^*(\theta)$ described in (17) is increasing, and when combined with the fee $F^*(\theta)$ in (18) it results in rents that are constant, at zero. Any bank with a proportion of bad loans θ will choose the corresponding contract $(s^*(\theta), F^*(\theta))$, foreclose or modify the amount $\gamma = \theta$, and obtain a transfer that just offsets the loss from foreclosing, meaning that this menu of two-part tariffs implements the contract in Proposition 1.

What are the fundamental properties of the model that mean that rents can be eliminated here? First, limited liability implies that participating banks will choose to foreclose or modify either none of their bad loans, or as many as they can, and second, the maximum amount of loans that they can foreclose or modify is the total amount of bad loans. We have used these two observations together to simplify the problem as highlighted in Remark 1. Third, limited liability also implies that the value of equity when not participating is convex in the proportion of bad loans θ. As we have illustrated, in the simplified problem, this means that the incentives to overstate the proportion of bad loans increase with θ. Since the incentive to understate can be increased by raising subsidies, both incentives can be played off against each other by offering a schedule of subsidies that increases with θ in a specific way.[14]

The optimal contract in Proposition 2 can also be implemented via a non-linear foreclosure or modification subsidy. Although maybe not the main focus of their paper, Aghion, Bolton, and Fries (1999) propose a scheme that can be interpreted as an alternative way of implementing our optimal contract based on a subsidy that is non-linear in the proportion of bad loans foreclosed. This can be translated into the terms of our model as follows: As in the two-part tariff, banks are allowed to privately choose the amount γ of loans that they want to foreclose or modify. Banks receive a subsidy $z(x)$ for foreclosing or modifying

[13]This is a special case of the argument of Maggi and Rodríguez-Clare (1995) who point out that, in general, decreasing convex outside opportunities can lead to optimal contracts that eliminate information rents for a range of agents. Remarkably, in our model this property holds globally due to the convexity of the participation value of equity in γ.

[14]The role of the same three properties of the model can also be observed in the proof of Proposition 1 in the appendix.

the additional, infinitesimal amount of bad loans dx, where $z(x)$ varies with the amount of foreclosed or modified loans as given by

$$\int_0^\gamma z(x)dx \equiv \Delta\pi_0(\gamma),$$

so that

$$z(x) = \frac{d\Delta\pi_0(x)}{dx}.$$

Since the subsidy associated with foreclosing or modifying an amount γ, $\int_0^\gamma z(x)dx$, is non-concave in γ, the value of equity when participating is still convex in γ, like in the two-part tariff case. Hence, banks would either foreclose or modify all bad loans, or no bad loans. But by construction, banks are again indifferent between foreclosing or modifying all bad loans or none. Under this subsidy, banks therefore participate, foreclose or modify all bad loans, and satisfy their participation constraint with equality. Hence, this is another way of implementing the optimal contract.

The optimal contract in Proposition 2 can also be implemented via a two-part tariff asset buyback like the one briefly described in the introduction. Suppose that a bank that reports a type θ^R commits to pay a fixed fee $F(\theta^R)$, in return for a price $p(\theta^R)$ per loan that it sells to the regulator. The regulator forecloses or modifies all loans that it buys. Following the argument above, the participation profits for a bank reporting type θ^R under this implementation are

$$\Pi(\theta, \theta^R) = 1 - \theta + p(\theta^R)\theta - D - F(\theta^R),$$

and the information rents of a bank that truthfully reports its type can be expressed as

$$U(\theta) = p(\theta^R)\theta - F(\theta^R) - (\pi_0(\theta) - (1 - \theta - D)).$$

The menu of two-part tariffs $\{p^*(\theta), F^*(\theta)\}$ under which any bank with a proportion θ will choose the right contract, sell all bad loans, and satisfy its participation constraint with equality, corresponding to the menu in (17) and (18), is given by

$$p^*(\theta) = 1 + \frac{d\pi_0(\theta)}{d\theta} < 1, \tag{19}$$

$$F^*(\theta) = -(\pi_0(\theta) - (1 - \theta - D)) + \theta p^*(\theta). \tag{20}$$

This implementation has as a main advantage over the two-part tariff foreclosure or modification subsidy that neither $p^*(\theta)$ nor $F^*(\theta)$ depend on ρ. For instance, as we show in the next section, if banks could foreclose or modify good loans and obtain a substantially higher recovery, a foreclosure subsidy could entice banks to overstate their proportion of bad loans.

Under an asset buyback scheme this situation cannot arise and the optimal contract can still be implemented.

Finally, we consider the question as to whether the optimal contract can be implemented via equity injections, in which participating banks receive an amount of cash T, in exchange for transferring a fraction λ of equity to the regulator. The regulator could offer a menu of such equity injections $\{\lambda(\theta), T(\theta)\}$. Unsurprisingly, the question as to whether or not this can implement the optimal contract hinges on whether the equity injections are *conditional* or *unconditional*, that is, whether a bank that chooses a particular equity injection from the menu is or is not required to foreclose a particular quantity of loans. Because our optimal contract makes transfers conditional on the quantity of loans that are foreclosed or modified, a scheme that aims to induce foreclosure or modification only through unconditional equity injections will not eliminate rents, and hence will be more costly than the type of schemes that we consider. This highlights an important point: If regulatory intervention is meant to affect not just bank solvency but also bank behaviour, in general, conditional schemes can be cheaper than unconditional schemes.

4 Foreclosing or modifying good loans

So far, we have assumed that banks can only foreclose, modify, or sell bad loans. However, if the type of loan is not verifiable even ex-post (even after the bank has indicated that a loan is a bad loan), banks might be able to foreclose,[15] modify, or sell good loans, in order to obtain higher transfers. In this section, we discuss to what extent our optimal contract is still incentive compatible when this is possible. We argue that an asset buyback implementation of our optimal contract is robust in this situation, and that a foreclosure or modification subsidy implementation is only robust as long as the recovery on good loans (that are foreclosed or modified) is not "substantially" higher than that on bad loans.

Suppose that foreclosing or modifying a good loan produces a recovery ρ_G, potentially different from the recovery obtained when foreclosing a bad loan, ρ. Assume that the amount recovered when foreclosing or modifying is non-verifiable.[16] For example, fundamentally

[15] Legally, there is only a basis for foreclosure if the terms of the loan contract have been breached. A breach of contract could be a default, or a violation of a covenant. It is plausible to interpret good loans as loans on which no default has occurred. However, covenants might have been violated. For example, loan contracts can stipulate that a firm maintains a minimum current ratio, defined as the ratio of current assets to current liabilities. If the current ratio falls below this level, the contract is breached. Such covenants are used in many loan contracts, are typically set very tight, and are hence frequently violated. Chava and Roberts (2008), for example, report that in their sample of loans to U.S. corporations between 1995 and 2005, about 15% of borrowers were violating covenants at any point in time, even when no default had occurred.

[16] If the amount recovered were verifiable, then the regulator could write contracts that only make payments

solvent borrowers, which are in violation of some covenant and on which the bank can therefore foreclose might have collateral that is worth ρ_G, even though a third party could not easily check the true value of this collateral. Alternatively, banks could modify loans of solvent borrowers by cutting face value substantially, but obtain other concessions from these borrowers, to obtain a combined value of ρ_G, but a third party could not easily find out about the other concessions the bank has obtained.

Suppose that $\rho_G < 1$, so that foreclosing or modifying a good loan is costly. In this case, banks would never foreclose or modify good loans in the absence of a scheme. This is because, conditional on survival, the change in the value of equity from foreclosing or modifying an additional good loan $\rho_G - 1$ is always negative. In contrast, conditional on survival, the change in the value of equity from foreclosing or modifying an additional bad loan $\rho - E[\varepsilon|\varepsilon > \bar{\varepsilon}]$ is positive if the bank is likely to survive, and negative if it is likely to fail, generating gambling incentives.

Consider a foreclosure or modification subsidy implementation of the optimal contract. If a bank is targeting a given transfer and therefore has to foreclose or modify a given quantity of loans, it will choose to foreclose or modify good loans instead of bad loans if the opportunity cost of doing so is lower. That is, if

$$\rho_G - 1 > \rho - E[\varepsilon|\varepsilon > \bar{\varepsilon}], \tag{21}$$

or

$$\rho_G - \rho > 1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}]. \tag{22}$$

We show in Appendix B that if $\rho_G - \rho$ is "large enough," banks may have incentives to foreclose or modify good loans to overstate their type and receive higher transfers. In this case, the subsidy implementation of our optimal contract would not be incentive compatible. Conversely, if $\rho_G - \rho$ is positive but "small enough," or non-positive, banks do not have incentives to overstate their type, and our optimal contract is incentive compatible.

The value of equity when foreclosing or modifying good loans is always increasing in the reported type (as long as the reported type receives a positive transfer under the scheme). This means that banks only consider foreclosing or modifying good loans if the value of equity they obtain from pretending to be of the "highest possible type" exceeds the value of equity from reporting truthfully. This determines the critical upper limit for $\rho_G - \rho$. The "highest possible type" might be determined by θ^*, the highest type that still receives a positive transfer in the baseline version of our scheme, or potentially by a technological limit on how many good loans a bank can pass off as bad loans (for example, a bank might only be legally able to foreclose on good loans that are in violation of a covenant).

to banks for bad loans.

It is worth noting that if the scheme is implemented as an asset buyback as discussed above, banks will never have incentives to overstate their type. Intuitively, this happens because under an asset buyback, the recovery when a loan is foreclosed or modified accrues to the regulator, and not to the bank. Therefore, even if $\rho_G > \rho$, the bank does not benefit from the higher recovery on the good loan when selling this instead of a bad loan, but the regulator does. Under a buyback implementation, banks therefore do not have incentives to sell good loans.

5 Imposing losses on debt holders

In the baseline version of our scheme, equity holders do not benefit from the scheme. However, debt becomes risk-free for the participating banks, implying that debt holders do benefit. This is a feature of almost any scheme that restores bank solvency. For this reason, there has been much debate about making debt holders contribute to the cost of bank rescues.[17] In this section we use a modified version of our scheme to explore to which extent a regulator can impose losses on debt holders, in a situation in which their consent is required. Although our discussion here is not intended to be a general treatment of this issue, it indicates that the ability to impose losses on debt holders is likely to crucially depend on the commitment power of the regulator.

Consider the following situation: Suppose that debt holders are atomistic and that they have the same information the regulator has. The regulator now not only offers a contract to the bank itself, but also to its debt holders. To simplify, we restrict ourselves to contracts for which the decision of debt holders is only whether to accept or reject an offer of the regulator. We also suppose that the timing is as follows: The regulator offers a contract to both debt holders and the bank. Debt holders decide first. On observing the decision of the debt holders, the regulator may then *revise* the offer made to the bank (but not to debt holders). We will consider the two extreme cases, in which the regulator either can or cannot commit ex-ante to not revising contracts, meaning that the regulator either can or cannot credibly threaten to punish debt holders if they do not accept an offer.

To be specific, the regulator offers banks a menu of contracts that specifies transfers to be received as a function of quantity of foreclosed loans, as before — in terms of the previous terminology, the regulator offers a schedule $T(\gamma)$. The contract also stipulates that debt holders must grant the regulator a call option on the debt with strike price $(1 - h)D$ that can be exercised when the bank participates, that is, if the bank chooses a contract from

[17] See, for example, Alan Greenspan's proposal mentioned in "Hire the A-Team," The Economist, August 7, 2008.

the menu $T(\gamma)$ under which it receives a non-zero transfer.[18] We will refer to the parameter $h \in [0,1]$ as the *haircut*, to be imposed on debt holders in the case that the bank chooses to participate.[19] The idea here is that debt holders of participating banks will be asked to contribute towards the cost of bailing out the bank.

It is important to note that the contract a bank picks from the menu $T(\gamma)$ does *not* depend on whether or not debt holders agree to grant the option, since from the point of view of the bank, it does not matter whether private parties or the regulator end up holding the debt.

We start with the case in which the regulator can commit and use the superscript C to denote parameter values specific to this case. Suppose the regulator announces a menu of contracts $T(\gamma)$ for which banks with $\theta \in [\hat{\theta}, \theta^C]$ will want to participate, but commits to only allowing a bank to participate when the bank's debt holders unanimously agree to a haircut h. What is the maximum haircut, h^C, that the regulator can impose in this case?

Let U_0^D denote the value of debt for a debt holder that does not accept the exchange offer. We can see that since the regulator has committed in this case to not letting the bank participate in the scheme, the value of debt becomes

$$U_0^D = D\Psi\left(\hat{\theta}\right) + D\int_{\hat{\theta}}^{1} R_0^D(\theta)\psi(\theta)d\theta, \tag{23}$$

where $R_0^D(\theta)$ is the expected fraction of face value recovered from a bank with bad assets θ when it is not bailed out,

$$R_0^D(\theta) = (1 - \Phi(\bar{\varepsilon}_0)) + \frac{1}{D}\int_0^{\bar{\varepsilon}_0} (1 - \theta + \theta\varepsilon)\phi(\varepsilon)d\varepsilon.$$

The value U_0^D accounts for the fact that debt holders obtain face value if the bank in question ends up having few bad loans ($\theta < \hat{\theta}$), and obtain an expected recovery otherwise. U_0^D also describes the payoff to a debt holder that does accept, when the required unanimity is not attained.

If all debt holders accept, the value of their debt (denoted as U^D) becomes

$$U^D(h, \theta^C) = D\Psi\left(\hat{\theta}\right) + (1-h)D\left[\Psi\left(\theta^C\right) - \Psi\left(\hat{\theta}\right)\right] + D\int_{\theta^C}^{1} R_0^D(\theta)\psi(\theta)d\theta. \tag{24}$$

The maximum haircut h^C that can be imposed is the one that makes debt holders just indifferent between accepting or rejecting and sets $U_0^D \equiv U^D(h, \theta^C)$. This is summarized in the next proposition:

[18] Equivalently, the regulator can offer to exchange the old debt claim for a new debt claim that is equivalent in all respects except that it includes the call option.

[19] Although it would be possible to condition the haircut on the θ revealed by the participating bank, doing so would not allow the regulator to extract additional rents, since debt holders are assumed to possess no private information.

Proposition 3. *When the regulator can commit, the optimal contract consists of the menu described in Corollary 1, with the set of participating banks equal to $\Theta_P = [\hat{\theta}, 1]$. Furthermore, the haircut h^C is such that $U^D(h^C, 1) = U_0^D$ or*

$$h^C = 1 - \int_{\hat{\theta}}^{1} R_0^D(\theta) \frac{\psi(\theta)}{1 - \Psi(\hat{\theta})} d\theta = 1 - E\left[R_0^D(\theta) \middle| \, \theta > \hat{\theta} \right].$$

The intuition for this result is quite straightforward. The decision of debt holders affects the bank only insofar as it might not be allowed to participate. If it is allowed to participate, the menu of contracts described in Corollary 1 still induces participating banks to foreclose or modify, and eliminates all information rents, such that bank equity holders are exactly as well off under the scheme as outside the scheme. Furthermore, under the scheme, all debt holders are exactly as well off as outside the scheme: the haircut is exactly equal to the expected losses if the bank is not allowed to participate.

The positive haircut reduces costs. Since the participation utilities of equity holders and debt holders are equal to their outside utilities, the regulator can appropriate any increase in net present value produced by the modification or foreclosure of bad loans. This makes the net cost of having any bank participate negative, and hence it is optimal to have all banks participate in the contract.

We now turn to the opposite case in which the regulator cannot commit at all. We use the superscript NC to denote parameter values specific to this case. The regulator announces a menu of contracts $T(\gamma)$ for which banks with $\theta \in [\hat{\theta}, \theta^{NC}]$ will want to participate, and states that only banks whose debt holders unanimously agree to a haircut h will be allowed to participate. However, the regulator now cannot commit to following through on this threat. What is the maximum haircut h^{NC} that the regulator can impose in this case?

Proposition 4. *When the regulator cannot commit, the optimal contract consists of the menu described in Corollary 1, with the set of participating banks equal to $\Theta_P = [\hat{\theta}, 1]$. Furthermore, the haircut h^C is such that $U^D(h^{NC}, 1) = U^D(0, \theta^*)$ or*

$$h^{NC} = \frac{1 - \Psi(\theta^*)}{1 - \Psi(\hat{\theta})} - \int_{\theta^*}^{1} R_0^D(\theta) \frac{\psi(\theta)}{1 - \Psi(\hat{\theta})} d\theta = \Pr(\theta > \theta^* | \theta > \hat{\theta}) \left(1 - E\left[R_0^D(\theta) \middle| \, \theta > \theta^* \right]\right)$$

The intuition for this result is similar to that for the preceding result. Again, the decision of debt holders affects the bank only insofar as it might not be allowed to participate. If it is allowed to participate, the menu of contracts described in Section 3 still induces participating banks to foreclose or modify, and eliminates all information rents, such that bank equity holders are exactly as well off under the scheme as outside the scheme.

However, debt holders now have a better outside option: If they refuse, in the second stage, the regulator will have to implement the optimal contract of the baseline version of our scheme (see Proposition 2) which means a zero haircut and that only banks with $\theta \in [\hat{\theta}, \theta^*]$ participate. This produces an expected value of debt which is lower than face value, but higher than the value of debt in the absence of intervention.

If the regulator wants to impose a positive haircut, it has to offer debt holders something in return. The only thing it can do, here, is to increase the set of banks that participate. It can do this such that the total expected transfers to debt holders remain constant. However, as more banks participate and foreclose or modify bad loans, this creates additional net present value, which the regulator can appropriate, since debt holders and equity holders are held to their outside option. This means that the cost of the scheme is decreasing in the number of banks that participate, and hence the regulator optimally has all banks participate. The haircut is then set exactly equal to the expectation of the losses that the debt holders would have faced for banks that would not have participated if the regulator had implemented the baseline scheme in the second stage.

As a corollary, the higher the social cost of funds, and hence the smaller the set of banks that participates if the regulator implements the baseline scheme in the second stage, the larger is the haircut that can be imposed on debt holders.

This argument highlights two points: First, in cases in which the consent of debt holders is necessary, one key to imposing losses on debt holders is likely to be the ability of the regulator to commit (that is, the ability to create a form of credible threat). Second, in such cases, if commitment is not possible (as is likely to be the case in practice), the limit to imposing losses on debt holders is in a sense determined by the ability of the regulator to fund a bail out when debt holders do not make concessions. Essentially, an inability to fund bailouts can be a form of commitment.

This suggests that in order to more easily impose losses, regulators should either look for ways of creating commitment, or find ways of relaxing the requirement of debt holder consent. Indeed, the current policy debate seems to revolve around the latter, as the discussion about contingent capital suggests (Flannery, 2009). The argument here also highlights that regardless of whether losses on debt holders can be imposed or not, information rents of equity holders can be eliminated via a version of our baseline scheme.

Finally, it is important to mention that in our model we treat debt holders as outside investors. Very often, however, many of the debt holders may themselves be banks or nonbank financial institutions. Thus, a haircut imposed on the debt of one bank might decrease the value of assets of another bank, and hence increase the need to bail out other banks.

6 Alternative welfare functions

In this section we discuss several variations of the social welfare function that we have used in the baseline model. We first consider how deposit insurance or social costs of bank failure could increase the attractiveness of getting banks to foreclose or modify their bad loans. We then discuss how crowding out effects on the one hand, and valuable long-run relationships between banks and their customers or negative externalities associated with foreclosures on the other hand could alter the attractiveness of getting banks to foreclose bad loans, as opposed to modifying them. Although these variations affect whether a regulator would prefer foreclosure to modification, and the set of banks that optimally participate in the scheme, they do not affect the mechanism design argument substantially, and the same type of contract can be used to eliminate information rents. Finally, as an example of a variation that complicates the mechanism design argument substantially, we consider how publicly observed participation decisions by banks might affect the possibility of bank runs.

As we pointed out in the previous section, in our baseline model bank debt holders benefit from the scheme. This is because bank debt becomes safe once banks stop gambling. The positive transfer that is necessary to induce banks to stop gambling is in fact an implicit transfer to debt holders. However, if the regulator already has some pre-existing commitments to make transfers to debt holders if a bank defaults (which can only happen when the bank gambles), then the *incremental* (expected) transfer to debt holders implied by the scheme over and above the expected transfers from pre-existing commitments, and hence the true *incremental* cost of the scheme, is lower.

Deposit insurance is such a pre-existing commitment to make transfers to (some) debt holders in the case of bank default. Suppose that insured deposits make up a fraction $\alpha \in [0, 1]$ of total bank debt D and that, for simplicity, α is the same across all banks. Assume that deposits are senior to other forms of debt, as is likely to be the case in practice, such that the regulator has to make insurance payments only if the remaining assets of a defaulting bank are less than αD. The expected deposit insurance cost associated with a bank with a proportion of bad loans θ that does not participate and decides to gamble is

$$DI(\theta) = \int_0^{\bar{\varepsilon}_{DI}} [\alpha D - (1 - \theta + \theta\varepsilon)]\phi(\varepsilon)d\varepsilon,$$

where $\bar{\varepsilon}_{DI}$ is the highest value of ε for which the remaining assets of the bank are not enough to repay αD. That is, $1 - \theta + \theta\bar{\varepsilon}_{DI} = \alpha D$. It is immediate that

$$DI'(\theta) = \int_0^{\bar{\varepsilon}_{DI}} (1 - \varepsilon)\phi(\varepsilon)d\varepsilon > 0,$$

$$DI''(\theta) = (1 - \bar{\varepsilon}_{DI})\phi(\bar{\varepsilon}_{DI})\frac{1 - \alpha D}{\theta^2} > 0,$$

so that the cost of deposit insurance increases in θ more than linearly.

The change in the incremental social cost of the scheme, which now becomes $\lambda(T(\theta) - DI(\theta))$, alters the cost-benefit balance. This means that in general it will be optimal to have a different (larger) set of banks participate. In particular, it is not necessarily true that the regulator will make only banks with a relatively low proportion of bad loans participate, and let banks with a high proportion of bad loans gamble, with the marginal type determined by an equation such as (11). As opposed what occurs in Figure 2, now the cost of making a bank participate, $\lambda\left(\Delta\pi_0(\theta) - DI(\theta)\right)$, is not necessarily convex in θ. Depending on the exact shape of $DI(\theta)$, which depends on α and the distribution of ε, it is possible, for example, that the regulator will make banks with low and high proportions of bad loans participate, but let those with medium proportions gamble. This could arise if the expected deposit insurance costs on banks with a medium proportion of bad loans were low, but the expected deposit insurance costs for banks with a high proportion of bad loans were high.

We now consider a situation in which bank failure might be costly *per se* from a social point of view. For simplicity, assume that there is a constant social cost $B > 0$ that is incurred whenever a bank fails. Now, making a bank participate and foreclose not only leads to an increase in social welfare derived from efficient foreclosure of the bad loans, $(\rho - E[\varepsilon])\theta$, but also to an increase in social welfare derived from the reduction of the expected social cost of bank failure from $B\Phi(\bar{\varepsilon}_0)$ to 0. The total social benefit of making a bank with type θ foreclose is now $(\rho - E[\varepsilon])\theta + B\Phi(\bar{\varepsilon}_0)$, and non-linear in θ. Depending on the distribution of ε, it it is again possible that the regulator will find it optimal, for example, to let banks with a low and high proportion of bad loans participate, but let those with medium proportions gamble. This could arise if, absent any intervention, the probability of bank failure for banks with a high proportion of bad loans is very high, so that the regulator will make such banks participate to ensure that they survive.[20]

There are also some arguments that specifically affect the attractiveness of foreclosures versus the attractiveness of modifications. On the one hand, foreclosures might have the additional benefit of facilitating creative destruction. Caballero, Hoshi, and Kashyap (2008) argue that zombie lending in Japan kept alive inefficient incumbents, and that this crowded out entry of more efficient firms, which in turn decreased growth. Although modifications can remove incentives for inefficient and insolvent borrowers to destroy value, they do not shut such borrowers down, and hence do not make room for more efficient entrants. Also, foreclosure might be preferable because the prospect of foreclosure can dissuade borrowers

[20] A more complicated version of the welfare function would arise if B, the social cost of bank failure, were to depend on the number (or type) of failing banks — as it plausibly might if the regulator is worried about an element of systemic risk. This would produce yet another set of banks that should optimally participate in the scheme.

from strategic default, whereas the prospect of modification cannot.

On the other hand, foreclosures might destroy long-run relationships between borrowers and their banks. These relationships can have a significant intrinsic value and they could be saved by modifying rather than foreclosing. Foreclosures might also produce negative externalities that could be avoided with modifications. For example, homes that are seized from residential real estate borrowers might be empty while the bank tries to find a new buyer, which might decrease the attractiveness of the whole neighbourhood. A large number of homes seized and sold might reduce property prices and hence recoveries to all banks.

There is therefore a question as to whether foreclosure is preferable to modification, or vice versa. Although this is an important question, a full answer is beyond the scope of this paper. We can conclude, however, that given a preferred action to take on bad loans (either foreclosure or modification), our model can accommodate the types of additional costs or benefits of this action as described above. Again, these are likely to change the set of banks that optimally participate.

Finally, the social cost of funds, λ, might not be constant but increasing in the total funds required for the scheme — in essence, assuming a linear cost of funds is an approximation that is reasonable in the context of small localized interventions, but it could be argued that it is not a reasonable for bailouts of the entire banking system. Obviously, this would affect the marginal bank that is rescued, but it would still be optimal to target the intervention at those banks that have a lower proportion of bad loans.

In any of the previous situations, the change in the costs and benefits from bailing out each bank affects the set of banks that optimally participates in the scheme. It is clear from Corollary 1 that in all of these cases, the same type of optimal contract can be used to eliminate information rents. There are, however, situations in which a variation in the social welfare function could complicate the mechanism design argument substantially. Consider a reduced form scenario in which banks are brought down by a bank run when their publicly perceived probability of default at $t = 2$, denoted as q, is above a certain threshold $\overline{q}$, and suppose that this produces a social cost.[21] In our context, the decision of a bank to participate reveals information about the probability of default of that bank. That is, if a bank participates, the probability of failure becomes 0. However, for banks that do not participate, depositors cannot distinguish whether this is because the bank is safe (in our context, $\theta \leq \hat{\theta}$), or because the bank is in such a dire condition that participation would be too costly (in our baseline context, $\theta > \theta^*$). As a result, the uninformed depositors would

[21] This could plausibly arise in a bank run model based on a global game (Goldstein and Pauzner, 2005), where the publicly perceived probability of default plays the role of the "fundamental."

calculate a probability of default conditional on a bank not participating, q_D, as

$$q_D = \frac{\int_{\theta^*}^{1} \Phi(\bar{\varepsilon}_0)\psi(\theta)d\theta}{1 - \Psi(\theta^*) + \Psi(\hat{\theta})},$$

where $\bar{\varepsilon}_0$ is obtained from (1). If it turns out that $q_D > \overline{q}$, banks that do not participate would be brought down by a bank run. The regulator can potentially prevent this situation and increase social welfare by changing the set of banks that do and do not participate in the scheme. This would give an additional criterion for selecting the banks that participate in the scheme. However, as in Philippon and Skreta (2012) or Tirole (2012), the non-participation value of equity would then be endogenous, which would further complicate the mechanism design problem.

7 Informational requirements of the scheme

We have shown that a scheme to prevent zombie lending can be designed to avoid information rents, even when the regulator does not know the true proportion of bad loans on banks' balance sheets. We have so far assumed, however, that the regulator knows other key quantities. First, in order to calculate the correct prices, subsidies, fees, or net transfers in the various different implementations, the regulator needs to know the loss from foreclosing, $\Delta\pi_0(\theta)$, which in turn is a function of the relevant leverage D, the recovery when taking immediate action on bad loans ρ, and the value of equity when banks do not participate $\pi_0(\theta)$. Second, in order to determine which banks should optimally participate, the regulator also needs to know the distribution of the proportion of bad loans θ in the population, $\Psi(\theta)$. In this section, we discuss to which extent these requirements can be circumvented by specific implementations, especially when banks have additional private information on some of the quantities.

While banks are probably unlikely to know much more about their leverage (D) than the regulator,[22] it is clear that they may be heterogeneous in their leverage. To the extent that these differences in leverage between banks are verifiable, however, a practical implementation of the scheme could simply condition on these: Since the "loss from foreclosing" can be seen to be increasing in leverage, banks with verifiably higher leverage should be offered menus with correspondingly larger transfers, that just compensate for the higher loss from foreclosing. Note, however, that due to the rent-elimination in our optimal contract, higher

[22] Although famously, Lehman Brothers tried to mask its leverage through "Repo 105" transactions, this type of transaction is open to legal challenge, as evidenced by the fraud charges filed against the auditors involved, and hence hopefully not very common.

transfers still do not mean positive rents.[23]

Banks are also heterogeneous with respect to the characteristics of the loans that they hold. These characteristics are likely to influence both ρ and $\phi(\varepsilon)$. To the extent that these characteristics are verifiable to the regulator, the scheme can condition on this. For instance, recoveries might differ for commercial real estate loans versus residential real estate loans, or for loans with different loan-to-value ratios, meaning that the regulator would have to adjust prices, subsidies, and net transfers in the various different implementations accordingly. Again, this would not affect rent elimination. However, it is possible that not all of the characteristics of loans are observable by the regulator. In this case, banks would have additional private information along a dimension other than just the quantity of bad loans.

These unobserved characteristics could affect the recovery when taking immediate action ρ. Since according to our scheme, banks with lower ρ need to receive higher net transfers, all banks would claim to have low ρ. That would constitute a problem for implementations via foreclosure or modification subsidies, particularly if ρ were not verifiable ex-post. However, it would not be a problem for asset buybacks. There, the prices paid and fees charged do not depend on ρ. A caveat is that it would not be clear which banks should participate, as the value from foreclosing that needs to be compared to the cost of foreclosure is $\rho - E[\varepsilon]$. In this case, although banks receive no rents from participating in the scheme, the regulator may fail to bail out some marginal banks or bail out some that should not be rescued. A similar problem exists if the regulator does not know the distribution of the proportion of bad loans in the population, $\psi(\theta)$.

The unobserved characteristics could also affect the distribution of recoveries when delaying action $\phi(\varepsilon)$, and hence the non-participation value of equity for gambling banks $\pi_0^G(\theta)$. Since according to our scheme, banks with better outside opportunities should get higher net transfers, all banks would claim to have distributions $\phi(\varepsilon)$ that imply better recoveries. Here, it is possible that combining our scheme with an auction design could help: In the asset buyback implementation, one can interpret the transaction in which a bank obtains the right to sell an unlimited quantity of bad loans at a given strike price in exchange for the participation fee as the purchase of a put option. Instead of selling these put options, they could be auctioned off. The idea is that banks would bid up the fees for the various options, and in doing so, reveal the value that they attach to the options and hence the value they attach to gambling, and about the distribution of ε. The actual design of such an auction scheme is likely to be non-trivial. For instance, it is not immediately obvious whether it

[23] This also means that banks cannot distort their leverage ex-ante to obtain a value of equity when participating that is different from the value of equity when not participating. In this sense, even a conditional version of the scheme does not distort ex-ante incentives. See also the short discussion of this issue in Section 3.2.

would be possible to ensure a sufficient degree of competition at such an auction.

Also, it would help if the distribution $\phi(\varepsilon)$ could be related to quantities that the regulator implicitly or explicitly has information about: First, $\phi(\varepsilon)$ is likely to be related to θ. For instance, if a bank gives credit to borrowers that do not receive credit elsewhere, it might end up not only with more bad loans, but also with bad loans that are backed by worse collateral. If knowledge of θ is sufficient to work out the correct $\phi(\varepsilon)$, it can be shown that, as long as the expressions for $\phi(\varepsilon)$ are replaced with the relevant $\phi(\varepsilon|\theta)$, the mechanism we describe will work in exactly the same way, as long as the non-participation value of equity is still decreasing and convex in the proportion of bad loans. The intuition is of course that the mechanism can already deal with private information about θ. Second, it is possible that $\phi(\varepsilon)$ might be related to ρ, because the quality of the collateral affects both ρ and $\phi(\varepsilon)$. In situations in which ρ is verifiable, the estimates of future recoveries $\phi(\varepsilon)$ used in the calculations could be conditioned on this, reducing the extra rents that a bank could extract from private information on the quality of collateral.

Interestingly, the kind of calculations that are necessary to work the out non-participation value of equity $\pi_0(\theta)$ are already performed by regulators when they undertake stress tests for the banks that they supervise. In these stress tests, regulators forecast losses for the banks under different macroeconomic scenarios. In order to forecast these losses, regulators need to forecast non-performing loan ratios, as well as future recoveries (θ and $\phi(\varepsilon)$ in terms of our model). The calculations for our mechanism would be substantially simpler and hence less error prone because regulators do not have to forecast non-performing loan ratios, as the information on these is extracted through the mechanism.

Even considering this as well as all the additional means of extracting information from banks as described above, it is possible that the regulator might make a mistake. We close by noting that a regulator could quantify the cost of making a mistake with one of the key inputs, using an explicit expression for welfare such as equation (W). Suppose the regulator believes that a specific alternative value for the input might be the true one. For this alternative value, and the given scheme being offered, the regulator could then calculate the set of participating banks, the resulting costs of the scheme, rents (if any), and the resulting difference in welfare.

8 Securitization

There are many countries in which most lending to the real economy is done via banks' balance sheets, and hence the kind of zombie lending by banks that we have described can be a significant problem, and our scheme might be useful.[24] However, in some countries, notably the USA, a substantial fraction of the lending to the real economy was and is financed via securitization. In this case, the loans are held by special purpose vehicles (SPVs), which sell (potentially tranched) securities to investors. Even when some such securities are bought by banks and hence the loans end up indirectly on banks' balance sheets, securitization can change the nature of the problem, because the collection of cash flows from borrowers (which includes decisions on whether to foreclose or modify loans) is delegated to a third party, the so-called servicer.

In a context in which securitization is important, a regulator might worry both about the incentives of banks that have large positions in "toxic" securities that have lost a substantial amount of value, and also directly about the incentives of servicers. Throughout the crisis, schemes have been proposed to tackle incentive problems at both levels.[25] To what extent can our proposed scheme provide useful ideas for dealing with problems arising from securitization?

Consider first the case of servicers. In many private-label commercial mortgage-backed securitization (CMBS) deals, servicers are also given an exposure to a first-loss tranche on the pool of mortgages. This essentially is a highly levered equity position. Gan and Mayer (2006) report that this is the case for about one third of the deals in their data. They interpret this as an attempt to align incentives of the servicer and the investors. They show that when delinquency rates are low, servicers that own a first-loss tranche put additional effort into efficiently managing bad mortgages, but that when delinquency rates are high, they slow the foreclosure process. These authors conclude that "servicers may be susceptible to the same kinds of problems that characterized undercapitalized banks when losses rose." In other words, our model applies in this type of case, with servicers taking the role of banks. An optimal scheme for this case could be implemented using an incentive payment to servicers whenever they modify or foreclose a mortgage, which varies with quantities as suggested by our model.

Servicers face very different incentives in residential mortgage-backed securitization (RMBS) deals as Levitin and Twomey (2011) describe.[26] They argue that given this contract, risk-

[24] In the introduction, we have discussed, for example, the case of Japan and the anecdotal evidence for some European countries such as Spain, which at the time of writing finds itself at the center of the Eurozone crisis precisely because of the bad state of its banking system.

[25] For example, the various actions proposed under the Troubled Asset Relief Program (TARP) act at the bank level, and the Home Affordable Modification Program (HAMP) acts at the servicer level.

[26] Key elements of the contract include: servicing fees which are a fixed percentage of remaining pool principal, ancillary fees from defaulting borrowers, the obligation to advance missed interest payments on delinquent loans to the pool (the advances are recovered if the loan is foreclosed, but no allowance is made for the interest cost of funding the advances), and the lack of compensation for the cost of modifications.

averse or liquidity constrained servicers are likely to have incentives to foreclose excessively, while servicers that have good access to liquidity might have incentives to excessively delay action on delinquent loans. Empirically, it appears to be the case that RMBS servicers have stronger incentives to foreclose than banks, as reported by Piskorski, Seru, and Vig (2010), and Agarwal, Amromin, Ben-David, Chomsisengphet, and Evanoff (2011). For this reason, the optimal incentive schemes for RMBS will probably look very different from the one that we propose for banks. We believe that the design of such schemes is likely to be an important topic for future research.

Now consider the case of banks that own the debt securities issued by Special Purpose Vehicles (SPVs) in securitization deals. The fundamental value of these securities is related to the incentives of the servicers to foreclose or modify the conditions of the individual loans. The incentives of servicers largely stem from the contracts they face. As the previous discussion suggests, many of these contracts unfortunately do not align the incentives of the servicers and the owners of the securities (a large fraction of which are banks). Renegotiation of such contracts can raise fundamental value. Unfortunately, due to co-ordination problems and conflicts of interests between the holders of different tranches, it can be difficult for all parties to agree on a new contract.[27] In theory, there would therefore be incentives for specialized investors to buy up all tranches of a deal to eliminate co-ordination problems and conflicts of interest, to be able to alter contracts, whenever this increases value.

However, just like in our model, the combination of losses on the securities, limited liability, and uncertainty about final payoffs of the toxic securities would imply that the less solvent banks would derive a "gambling value" from the uncertainty about payoffs of the toxic securities, over and above the fundamental value, and hence might not be prepared to sell at a reasonable price. Clearly, the private sector will not be prepared to lose money in order to solve the problem, and hence regulatory intervention might be called for.[28] For example, a regulator might implement an asset buyback version of our scheme to buy all tranches of toxic securities (at high prices that reflect the "gambling value"), and then sell them on to outside investors (at lower prices that reflect fundamental value). The outside investors could then renegotiate the contracts of servicers in order to maximize value. The main difference with the optimal contract described in Section 3 is the added difficulty of ensuring that *all*

[27] For an illustrative example of the problems when several parties hold different tranches, see e.g. "Hancock at Center of 'Tranche Warfare"', The Wall Street Journal, January 21, 2009.

[28] In practice, the main concern seems to have been that outside investors were financially constrained. The Public-Private Investment Program for Legacy Assets (part of the Troubled Asset Relief Program) provided subsidized loans to outside investors for the purchase of and legacy securities and legacy loans. The fact that to date only a fraction of the earmarked funds have been lent could suggest that the "gambling value" effect might be preventing outside investors from buying tranches at reasonable prices. See the documentation on `http://www.treasury.gov/initiatives/financial-stability/programs`

securities issued by any given SPV are bought, in order to allow renegotiation. This might mean, for instance, that in order to obtain all relevant securities, the regulator essentially has to bail out all banks. For this policy to be welfare-enhancing, the social cost of funds would have to be low.

An additional (and related) problem is that banks might prefer to hang on to the toxic securities instead of engaging in new (safe, NPV positive) lending — a form of debt overhang. The regulator might want to remove the toxic securities from balance sheets to eliminate this debt overhang, even if altering the contracts of servicers does not affect fundamental value. Suppose that, again, banks have a proportion θ of toxic securities on their balance sheet. Suppose now that banks can give safe new loans that produce a net return of $r > 0$, but that they cannot easily expand their balance sheet. This might occur because banks are capital constrained and cannot easily raise equity, or because they cannot easily raise debt financing, which is plausibly the case if there is private information on θ and hence there is some form of adverse selection in the markets for funding. But suppose that banks know that they could sell the toxic securities for cash, at a price $\rho' = E[\varepsilon]$ reflecting fundamental value, and then lend out the cash to obtain a net rate of return of r. The tradeoff here is losing the random ε, but gaining a certain $\rho := \rho'(1+r) > E[\varepsilon]$. It is clear that at this point, our model applies: Banks with $\theta > \hat{\theta}$ would hang on to the toxic securities to avoid revealing their losses, and not engage in new lending. In this context, an asset buyback along the lines that we describe would help to clear out toxic assets from balance sheets and stimulate new lending, even though this per se would not get bad loans renegotiated or foreclosed.[29]

9 Concluding remarks

Banks that are insolvent but still operating have incentives to avoid the crystallization of losses on their bad loans, to hide the fact that they are insolvent and gamble for resurrection. This can take the form of banks deciding to roll over loans to insolvent borrowers (sometimes referred to as "zombie lending"), or refusing to modify the terms of loans in favor of the borrower even when this is clearly necessary.

We consider how a welfare-maximizing regulator would optimally deal with this zombie lending problem, even when banks can hide bad loans by avoiding the crystallization of losses, and hence have private private information on the true proportion of bad loans on their balance sheet. When designing schemes to deal with this problem, it is important to try to minimize the information rents to equity holders that might arise from the private

[29] In a situation in which the ability of banks to grow their balance sheet is related to the degree of adverse selection in markets, a fuller discussion of debt overhang and asset buybacks would also take into account mechanism-dependent participation constraints (Tirole, 2012; Philippon and Skreta, 2012).

information.

We have proposed a voluntary scheme that can either be interpreted as a form of asset buyback in which the regulator buys bad loans from banks and then forecloses or modifies, or as a scheme that subsidizes the foreclosure or modification of bad loans by the banks. The key feature of the scheme is that it uses a form of price discrimination, which can reduce rents. We show that in the context we consider, price discrimination cannot just reduce information rents to equity holders, but can actually completely eliminate them. Importantly, we show how the fundamental features of the problem that can cause the zombie lending in the first place, namely limited liability of banks and the risk inherent in hanging on to bad loans, are closely related to the features of the problem that allow the elimination of rents to equity holders.

The paper suggests avenues for future research. For example, we have assumed that regulators maximize welfare, and hence wants to address the zombie lending problem. In practice, in the same way that banks engage in forbearance lending in order to hide their financial situation, regulators can sometimes have incentives to engage in regulatory forbearance: They may prefer not to act on the fact that some banks are insolvent, hoping that the economic situation recovers, that banks become solvent again, and that it is never revealed that they had failed to identify problems to start with. In this paper, we have left these political economy issues aside in order to be able to focus on identifying the optimal scheme. However, understanding the incentives of regulators in the context of this specific problem is also likely to be of practical importance, and warrants more research.

Also, the type of rent-eliminating price discrimination mechanism that we identify here is likely to have applications beyond the prevention of zombie lending. As briefly mentioned in Section 8, this mechanism could be applied, for instance, to the indirect stimulation of lending, and help in solving debt overhang problems. However, it could also be applied to direct stimulation of lending, by informing the design of lending subsidy schemes:[30] One potential problem with subsidizing bank lending is that banks can always easily give a large number of risky new loans with negative net present value, but probably could only give a limited number of safe new loans with positive net present value. Since banks know more about the true number of safe new loan opportunities than the regulator, subsidy schemes run the danger of inducing banks to give only risky loans, as a form of risk-shifting. We conjecture that a mechanism that is technically similar to the one that we describe in the paper could be used to subsidize banks to only give safe new loans with positive net present

[30] There have been schemes of this type, see e.g. the recent Funding for Lending Scheme of the Bank of England, launched July 13, 2012, as described here: `http://www.bankofengland.co.uk/markets/Pages/FLS/default.aspx`

value, again, without providing information rents.

References

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, and D. D. Evanoff, 2011, "The role of securitization in mortgage renegotiation," *Journal of Financial Economics*, 102, 559–578.

Aghion, P., P. Bolton, and S. Fries, 1999, "Optimal Design of Bank Bailouts: The Case of Transition Economies," *Journal of Institutional and Theoretical Economics*, 155, 51–70.

Ausubel, L. M., and P. Cramton, 2008, "A Troubled Asset Reverse Auction," http://www.ausubel.com/auction-papers/ausubel-cramton-troubled-asset-reverse-auction.pdf.

Baron, D., and R. Myerson, 1982, "Regulating a Monopolist with Unknown Cost," *Econometrica*, 50, 911–930.

Bhattacharya, S., and K. G. Nyborg, 2010, "Bank Bailout Menus," Swiss Finance Institute Working Paper 10-24.

Bolton, P., and D. S. Scharfstein, 1996, "Optimal Debt Structure and the Number of Creditors," *Journal of Political Economy*, 104, 1–25.

Caballero, R. J., T. Hoshi, and A. K. Kashyap, 2008, "Zombie Lending and Depressed Restructuring in Japan," *American Economic Review*, 98, 1943–1977.

Chava, S., and M. R. Roberts, 2008, "How Does Financing Impact Investment? The Role of Debt Covenants," *Journal of Finance*, 63, 2085–2121.

Corbett, J., and J. Mitchell, 2000, "Banking Crises and Bank Rescues: The Effect of Reputation," *Journal of Money, Credit, and Banking*, 32, 474–512.

Diamond, D. W., and R. G. Rajan, 2011, "Fear of fire sales, illiquidity seeking, and credit freezes," *Quarterly Journal of Economics*, 126, 557–591.

Flannery, M. J., 2009, "Stabilizing Large Financial Institutions with Contingent Capital Certificates," Working Paper.

Fudenberg, D., and J. Tirole, 1991, *Game Theory*, MIT Press, Cambridge MA USA.

Gan, Y. H., and C. Mayer, 2006, "Agency Conflicts, Asset Substitution, and Securitization," NBER Working Paper 12359.

Giammarino, R., 1989, "The Resolution of Financial Distress," *Review of Financial Studies*, 2, 25–47.

Giannetti, M., and A. Simonov, forthcoming, "On the Real Effects of Bank Bailouts: Micro-Evidence from Japan," *American Economic Journal: Macroeconomics*.

Goldstein, I., and A. Pauzner, 2005, "Demand-Deposit Contracts and the Probability of Bank Runs," *Journal of Finance*, 60, 1293–1327.

Gorton, G., and L. Huang, 2004, "Liquidity, Efficiency and Bank Bailouts," *American Economic Review*, 94, 455–483.

Haugen, R. A., and L. W. Senbet, 1978, "The Insignificance of Bankruptcy Costs to the Theory of Optimal Capital Structure," *Journal of Finance*, 33, 383–393.

House, C. L., and Y. Masatlioglu, 2010, "Managing Markets for Toxic Assets," NBER Working Paper No. 16145.

Jullien, B., 2000, "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93, 1–47.

Klemperer, P., 2010, "The Product-Mix Auction: a New Auction Design for Differentiated Goods," *Journal of the European Economic Association*, 8, 526–536.

Levitin, A. J., and T. Twomey, 2011, "Mortgage Servicing," *Yale Journal on Regulation*, 28.

Lewis, T., and D. Sappington, 1989, "Countervailing Incentives in Agency Models," *Journal of Economic Theory*, 49, 294–313.

Maggi, G., and A. Rodríguez-Clare, 1995, "On Countervailing Incentives," *Journal of Economic Theory*, 66, 238–263.

Mitchell, J., 1998, "The Problem of Bad Debts: Cleaning Banks' Balance Sheets in Economies in Transition," CEPR Discussion Paper No. 1977.

Mitchell, J., 2001, "Bad Debts and the Cleaning of Banks' Balance Sheets: An Application to Transition Economies," *Journal of Financial Intermediation*, 10, 1–27.

Peek, J., and E. S. Rosengren, 2005, "Unnatural Selection: Perverse Incentives and the Misallocation of Credit in Japan," *American Economic Review*, 95, 1144–1166.

Philippon, T., and P. Schnabl, forthcoming, "Efficient Recapitalization," *Journal of Finance*.

Philippon, T., and V. Skreta, 2012, "Optimal Interventions in Markets with Adverse Selection," *American Economic Review*, 102, 1–28.

Piskorski, T., A. Seru, and V. Vig, 2010, "Securitization and distressed loan renegotation: Evidence from the subprime mortgage crisis," *Journal of Financial Economics*, 97, 369–397.

Rogerson, W. P., 1987, "On the Optimality of Menus of Linear Contracts," Northwestern University CMS-EMS Discussion Paper No. 714R.

Sekine, T., K. Kobayashi, and Y. Saita, 2003, "Forebearance Lending: The Case of Japanese Firms," *Monetary and Economic Studies*, pp. 69–92.

Tirole, J., 2012, "Overcoming Adverse Selection: How Public Intervention Can Restore Market Functioning," *American Economic Review*, 102, 29–59.

Appendix

A Proofs

Proof of Lemma 1: We first show that the value of equity is convex in γ: We note that the derivative of (2) with respect to γ is given by

$$\int_{\bar{\varepsilon}_0(\gamma)}^{1} (\rho - \varepsilon)\phi(\varepsilon)d\varepsilon, \tag{25}$$

and that the second derivative becomes

$$-(\rho - \bar{\varepsilon}_0(\gamma))\phi(\bar{\varepsilon}_0)\frac{\partial \bar{\varepsilon}_0}{\partial \gamma}. \tag{26}$$

To evaluate the sign of the second derivative, it is useful to note that

$$\rho - \bar{\varepsilon}_0(\gamma) = \frac{(1-D)-(1-\rho)\theta}{\theta - \gamma} = -\frac{\partial \bar{\varepsilon}_0}{\partial \gamma}(\theta - \gamma). \tag{27}$$

Consider first banks for which $\theta = \frac{1-D}{1-\rho} \equiv \bar{\theta}$. For such banks, $\bar{\varepsilon}_0 = \rho$, regardless of γ, and hence the second derivative is always zero. Checking (25), however, we can see that for such banks, the first derivative will always be negative, and hence such banks will foreclose the minimum amount $\gamma = 0$ and gamble.

Consider now banks for which $\theta \neq \bar{\theta}$. For such banks, as indicated by (27), $\rho - \bar{\varepsilon}_0$ and $\partial \bar{\varepsilon}_0 / \partial \gamma$ have always the opposite sign, and since $\phi(\varepsilon) > 0$, the second derivative is positive. As a result, the value of equity is convex in γ, and the optimal choice of γ is either 0 or θ.

Furthermore, note that $\pi_0^F(0) = \pi_0^G(0)$, that $\pi_0^F(\theta) = 0$ and $\pi_0^G(\theta) > 0$ for $\theta > \bar{\theta}$, that $\pi_0^F(\theta)$ is continuous, decreasing, and linear in θ, that $\pi^G(\theta)$ is continuous, decreasing, and convex in θ and that

$$\left.\frac{d\pi_0^G(x)}{dx}\right|_{x=0} = -(1 - E[\varepsilon]) < -(1-\rho) = \left.\frac{d\pi_0^F(x)}{dx}\right|_{x=0} \tag{28}$$

since $E[\varepsilon] < \rho$. It follows that there exists a unique $\hat{\theta} > 0$ such that for $0 < \theta < \hat{\theta}$, $\pi_0^G(\theta) < \pi_0^F(\theta)$, and for $\theta > \hat{\theta}$, $\pi_0^G(\theta) > \pi_0^F(\theta)$. Since $\pi_0^F(\bar{\theta}) = 0$, this also implies that $\hat{\theta} < \bar{\theta}$.

We can see that banks with $\theta < \hat{\theta}$ foreclose, and that banks with $\theta > \hat{\theta}$ gamble. (In the paper, we focus on the case in which $\hat{\theta} < 1$, such that the set of banks that gamble is not empty.) $\square$

Proof of Proposition 1: It is obvious that the proposed contract satisfies (PC) with equality. It remains to be shown that it satisfies (IC). In order to do so, we need to show that a bank maximizes its participation value of equity Π when reporting $\theta^R = \theta$. We note that the standard (first-order) approach cannot be used to show incentive compatibility, because for the contract in question, the first derivative of Π with respect to θ^R is discontinuous at $\theta^R = \theta$. Instead, we rely on three key properties of the model to establish that Π is maximized at $\theta^R = \theta$: The first property is that the first derivative of Π with respect to γ is not defined at $\gamma = 0$ and $\gamma = \theta$. The second property is that Π is a convex function of γ and T. The third property is that the non-participation value of equity $\pi_0(\theta)$ is convex in θ, which implies that the transfer $T(\theta^R)$ that is specified in the contract will be convex in θ^R.

The first property is a consequence of the assumption that banks can only foreclose bad loans (or that there is a wedge in payoffs between good and bad loans in the case in which they can also foreclose good loans, see Section 4). Both the second and third property are a consequence of limited liability, and our assumptions on how γ, T, and θ relate to asset payoffs.

Due to the first property, for the contract under consideration (for which $\gamma(\theta^R) = \theta^R$), we can immediately establish that $\theta^R = 0$ and $\theta^R = \theta$ are critical points at which local maxima in Π can occur.

We now use the second and third property to show that for the contract under consideration, Π is convex in θ^R, for $\theta^R \in (0, \theta)$. Since θ^R only affects Π through γ and T, we can write the second derivative of $\Pi(\gamma(\theta^R), T(\theta^R), \theta)$ with respect to θ^R as

$$\frac{d^2\Pi}{(d\theta^R)^2} = \underbrace{\frac{\partial^2\Pi}{\partial\gamma^2}\left(\frac{d\gamma}{d\theta^R}\right)^2 + 2\frac{\partial^2\Pi}{\partial\gamma\partial T}\frac{d\gamma}{d\theta^R}\frac{dT}{d\theta^R} + \frac{\partial^2\Pi}{\partial T^2}\left(\frac{dT}{\theta^R}\right)^2}_{\text{Term I}} \underbrace{+\frac{\partial\Pi}{\partial\gamma}\frac{d^2\gamma}{d(\theta^R)^2}}_{\text{Term II}} \underbrace{+\frac{\partial\Pi}{\partial T}\frac{d^2T}{d(\theta^R)^2}}_{\text{Term III}} \quad (29)$$

Define $z = (d\gamma/d\theta^R, dT/d\theta^R)'$, and let H be the Hessian of Π w.r.t. (γ, T). Then Term I can also be written as the quadratic form $z'Hz$. Due to the second property, we know that H is positive semi-definite,[31] implying that the quadratic form is non-negative, regardless of the contract $(\gamma(\theta^R), T(\theta^R))$. The signs of Term II and Term III depend on the curvature of the contract. For the contract under consideration, Term II is zero, and Term III is non-negative because of the third property. Hence we have established that for the contract under consideration, $\frac{d^2\Pi}{(\theta^R)^2} \geq 0$ for $\theta^R \in (0, \theta)$.

The convexity of Π w.r.t θ^R in the range $(0, \theta)$ implies that a global maximum of Π w.r.t. θ^R is either at $\theta^R = 0$, or at $\theta^R = \theta$, or at both points (and that we can ignore any potential stationary point in $(0, \theta^R)$ because it will not be a maximum). We note that (i) $\Pi(\theta, 0) = \pi_0^G(\theta)$, i.e. that reporting a type of 0 produces the same value of equity as when

[31] More specifically, the Hessian implied by (6) has one positive eigenvalue, and one zero eigenvalue.

not participating and gambling, and that (ii) $\Pi(\theta, \theta) = \pi_0(\theta)$ by the definition of $\Delta\pi_0(\theta)$, i.e. that reporting truthfully produces the same equity value as when not participating and taking the privately optimal action (either foreclosing or gambling).

Banks with type θ such that $\theta \leq \hat{\theta}$ want to foreclose outside the scheme, since for them, $\pi_0(\theta) = \pi_0^F(\theta) \geq \pi_0^G(\theta)$. Here, this (trivially) means that $\Pi(\theta, 0) = \pi_0^G(\theta) \leq \pi_0(\theta) = \Pi(\theta, \theta)$ and therefore it is optimal for them to report their type truthfully. Banks with a type θ such that $\theta > \hat{\theta}$ want to gamble outside the scheme, since for them, $\pi_0(\theta) = \pi_0^G(\theta)$ $(\geq \pi_0^F(\theta))$. By construction, $\Pi(\theta, \theta) = \pi_0^G(\theta)$ for such banks and they are therefore indifferent between truthfully reporting their type or lying and reporting a type of $\theta^R = 0$. Together, this implies that the proposed contract must be incentive compatible. □

Proof of Proposition 2: It is obvious that the proposed optimal contract satisfies (PC) with equality. From direct consideration of the welfare function it is immediate that the proposed contract maximizes welfare, subject to the constraint (PC) (see the main text for a verbal argument). Apply Corollary 1 to see that the proposed contract is incentive compatible. □

B Foreclosing or modifying good loans

As explained in Section 4, we suppose that foreclosing or modifying a good loan produces a recovery $\rho_G < 1$, potentially different from the recovery obtained when foreclosing or modifying a bad loan, ρ. The amount recovered is unverifiable so that a regulator cannot contract on this. We consider the foreclosure subsidy implementation, and the asset buyback implementation of our optimal contract. (The case of the modification subsidy implementation is exactly analogous.) We show that the foreclosure subsidy implementation of the contract in Proposition 2 is incentive compatible as long as $\rho_G - \rho$ is "small enough" (in a sense to be made precise below), and that the asset buyback implementation is always incentive compatible.

Foreclosure subsidy implementation Consider a foreclosure subsidy implementation of the optimal contract (the case of the modification subsidy is analogous). If a bank is targeting a given transfer and therefore has to foreclose a given amount of loans, it will foreclose good loans if the opportunity cost of doing so is lower than the cost of foreclosing bad loans. That is, if

$$\rho - E[\varepsilon | \varepsilon > \bar{\varepsilon}] < \rho_G - 1, \tag{30}$$

or

$$\rho_G - \rho > 1 - E[\varepsilon | \varepsilon > \bar{\varepsilon}]. \tag{31}$$

As long as $\rho_G > \rho$, it is possible that some banks that are very unlikely to survive (and hence have a high $E[\varepsilon|\varepsilon > \bar{\varepsilon}]$) foreclose good loans before foreclosing bad loans. This happens when the probability of survival is so small ($\bar{\varepsilon}$ is so high) that the expected return conditional on survival of bad loans that are rolled over is very similar to the return on good loans, and the recovery on good loans is much higher than the recovery on bad loans. Since foreclosing good loans makes a bank even less likely to survive (increases $\bar{\varepsilon}$ and hence $E[\varepsilon|\varepsilon > \bar{\varepsilon}]$), a bank that starts foreclosing good loans would foreclose all good loans before considering foreclosing bad loans. If $\rho_G \leq \rho$, all banks will always foreclose all bad loans before considering foreclosing good loans.

We first consider the case where $\rho_G \leq \rho$, and then consider the case where $\rho_G > \rho$.

Case I: $\rho_G \leq \rho$. In this case, banks will only consider foreclosing good loans once they have already foreclosed all bad loans. Under our optimal contract, if a bank reports type θ^R, where $\theta^R > \theta$, it will therefore have to foreclose an amount $\theta^R - \theta$ of good loans in addition to foreclosing all of its bad loans. Its value of equity would then be

$$\Pi(\theta, \theta^R) = \underbrace{1 - \theta - (\theta^R - \theta)}_{\text{remaining good loans}} + \underbrace{(\theta^R - \theta)\rho_G}_{\text{foreclosed good loans}} + \underbrace{\theta\rho}_{\text{foreclosed bad loans}} - D + \Delta\pi_0(\theta^R) \tag{32}$$

or, rearranging and inserting the expression for $\Delta\pi_0(\theta^R)$,

$$\Pi(\theta, \theta^R) = \pi_0(\theta^R) + (\rho_G - \rho)(\theta^R - \theta) \leq \pi_0(\theta^R) < \pi_0(\theta). \tag{33}$$

Since the value of equity from participating and truthfully reporting is equal to $\pi_0(\theta)$, a bank would therefore never have incentives to overreport its type, and the optimal contract is robust in this case.

Case II: $\rho_G > \rho$. Here, we need to distinguish two subcases. Define the proportion of bad loans $\theta^\dagger$ as the proportion for which

$$1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}] = \rho_G - \rho. \tag{34}$$

Since $\bar{\varepsilon}$ and hence $E[\varepsilon|\varepsilon > \bar{\varepsilon}]$ are increasing in θ, banks with $\theta < \theta^\dagger$ are so safe that for them, foreclosing bad loans is less costly than foreclosing good loans. Since foreclosing some bad loans makes them safer, they will foreclose all bad loans before foreclosing any good loans. Conversely, banks with $\theta > \theta^\dagger$ will be so risky that for them, foreclosing bad loans will be more costly than foreclosing good loans. Since foreclosing some good loans makes them even riskier, they will foreclose all good loans before foreclosing any bad loans.

Consider first the safer banks for which $\theta < \theta^\dagger$. Using the same argument as in the previous case, we can work out that such banks, when reporting θ^R, have a value of equity of

$$\Pi(\theta, \theta^R) = \pi_0(\theta^R) + (\rho_G - \rho)(\theta^R - \theta). \tag{35}$$

We note that this expression is convex in θ^R, which implies that either, banks will want to report truthfully, or overstate their type as much as possible. Since the highest type that still obtains a transfer is θ^*, we see that such banks will not want to overstate their type at all as long as the recovery ρ_G on good loans is not much larger than the recovery on bad loans ρ, or

$$\pi_0(\theta^*) + (\rho_G - \rho)(\theta^* - \theta) < \pi_0(\theta), \tag{36}$$

which can be rewritten as

$$\rho_G - \rho < \frac{\pi_0(\theta) - \pi_0(\theta^*)}{\theta^* - \theta}. \tag{37}$$

Consider now the riskier banks for which $\theta > \theta^\dagger$. We separately consider the case in which $\theta^R < 1 - \theta$, i.e. banks that foreclose some of their good loans but none of their bad loans, and the case in which $\theta^R > 1 - \theta$, in which banks foreclose all of their good loans and some of their bad loans.

When $\theta^R < 1 - \theta$, banks foreclose an amount θ^R of their good loans, and none of their bad loans. We can write the value of equity as

$$\Pi(\theta, \theta^R) = \int_{\bar{\varepsilon}}^1 \left(\underbrace{1 - \theta - \theta^R}_{\text{remaining good loans}} + \underbrace{\theta^R \rho_G}_{\text{foreclosed good loans}} + \underbrace{\theta \varepsilon}_{\text{remaining bad loans}} - D + \Delta\pi_0(\theta^R) \right) \phi(\varepsilon) d\varepsilon, \tag{38}$$

for a suitably defined $\bar{\varepsilon}$.

Rearranging and inserting the expression for $\Delta\pi_0(\theta^R)$, we obtain

$$\Pi(\theta, \theta^R) = \int_{\bar{\varepsilon}}^1 \left(\theta^R(\rho_G - \rho) - \theta(1 - \varepsilon) + \pi(\theta^R)\right) \phi(\varepsilon) d\varepsilon. \tag{39}$$

There is now a tradeoff: Foreclosing good loans means a higher recovery of (term in $\rho_G - \rho$), but also means exchanging the return on good loans against the return on bad loans (term in $1 - \varepsilon$).

Taking derivatives with respect to θ^R, we can see that

$$\frac{\partial \Pi(\theta, \theta^R)}{\partial \theta^R} = \int_{\bar{\varepsilon}}^1 \left((\rho_G - \rho) + \frac{d\pi_0(\theta^R)}{d\theta^R} \right) \phi(\varepsilon) d\varepsilon = (1 - \Phi(\bar{\varepsilon})) \left((\rho_G - \rho) + \frac{d\pi_0(\theta^R)}{d\theta^R} \right), \tag{40}$$

which is positive iff $\rho_G - \rho > -\frac{d\pi_0(\theta^R)}{d\theta^R}$. But since

$$-\frac{d\pi_0(\theta^R)}{d\theta^R} = \int_{1-(1-D)/\theta^R}^1 (1 - \varepsilon)\phi(\varepsilon) d\varepsilon = (1 - \Phi(\bar{\varepsilon}))(1 - E[\varepsilon | \varepsilon > \bar{\varepsilon}]), \tag{41}$$

we can see that

$$\rho_G - \rho > 1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}] > (1 - \Phi(\bar{\varepsilon}))(1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}]), \tag{42}$$

i.e. this derivative is always positive. This means that such banks will foreclose as many of their good loans as possible.

When $\theta^R > 1-\theta$, banks foreclose all of their good loans, $1-\theta$, and an amount $\theta^R-(1-\theta)$ of bad loans. In other words, they roll over an amount $\theta - (\theta^R - (1-\theta)) = 1 - \theta^R$ of bad loans. We can write the value of equity as

$$\Pi(\theta, \theta^R) = \int_{\bar{\varepsilon}}^{1} \Bigg(\underbrace{(1-\theta)\rho_G}_{\text{foreclosed good loans}} + \underbrace{(\theta^R - (1-\theta))\rho}_{\text{foreclosed bad loans}} + \underbrace{(1-\theta^R)\varepsilon}_{\text{remaining bad loans}} - D + \Delta\pi_0(\theta^R) \Bigg) \phi(\varepsilon) d\varepsilon \tag{43}$$

Rearranging and inserting the expression for $\Delta\pi_0(\theta^R)$, we obtain

$$\Pi(\theta, \theta^R) = \int_{\bar{\varepsilon}}^{1} \left((1-\theta)(\rho_G - \rho) - (1-\theta^R)(1-\varepsilon) + \pi_0(\theta^R) \right) \phi(\varepsilon) d\varepsilon. \tag{44}$$

Taking derivatives with respect to θ^R, we can see that

$$\begin{aligned}
\frac{\partial \Pi(\theta, \theta^R)}{\partial \theta^R} &= \int_{\bar{\varepsilon}}^{1} \left(1 - \varepsilon + \frac{d\pi_0(\theta^R)}{d\theta^R} \right) \phi(\varepsilon) d\varepsilon && (45)\\
&= (1 - \Phi(\bar{\varepsilon})) \left((1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}) - (1 - \Phi(\bar{\varepsilon}))(1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}]) \right) && (46)\\
&= (1 - \Phi(\bar{\varepsilon}))\Phi(\bar{\varepsilon})(1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}]) > 0, && (47)
\end{aligned}$$

i.e. this derivative is always positive. This means that such banks will want to overstate their type as much as is possible.

Since the highest type that still obtains a transfer is θ^R, we see that banks with $\theta > \theta^\dagger$ will not want to overstate their type as long as

$$\int_{\bar{\varepsilon}}^{1} \left(\min(\theta^*, 1-\theta)(\rho_G - \rho) - \min(\theta, 1-\theta^*)(1-\varepsilon) + \pi_0(\theta^*) \right) \phi(\varepsilon) d\varepsilon < \pi_0(\theta), \tag{48}$$

$$(1 - \Phi(\bar{\varepsilon})) \left(\min(\theta^*, 1-\theta)(\rho_G - \rho) - \min(\theta, 1-\theta^*)(1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}]) + \pi_0(\theta^*) \right) < \pi_0(\theta), \tag{49}$$

or

$$\rho_G - \rho < \frac{1}{\min(\theta^*, 1-\theta)} \left(\frac{\pi_0(\theta)}{1 - \Phi(\bar{\varepsilon})} - \pi_0(\theta^*) + \min(\theta, 1-\theta^*)(1 - E[\varepsilon|\varepsilon > \bar{\varepsilon}) \right), \tag{50}$$

for suitably defined $\bar{\varepsilon}$. We note that the right-hand side of the previous expression is always bigger than zero.

We can see that in general, when $\rho_G \geq \rho$, as long as the difference $\rho_G - \rho$ is small enough, banks will not have incentives to overstate their type.

Asset buyback implementation If the scheme is implemented as an asset buyback as discussed in Section 3.3, banks will never have incentives to overstate their type. Intuitively, this happens because under an asset buyback, the recovery when a loan is foreclosed accrues to the regulator, and not to the bank. Therefore, even if $\rho_G > \rho$, the bank does not benefit from the higher recovery on the good loan when selling this instead of a bad loan, but the regulator does. Under a buyback implementation, banks therefore never have incentives to sell good loans to obtain higher transfers. (We skip the formal argument here, but note that it is similar to the foreclosure subsidy argument for Case I: $\rho_G \leq \rho$ above.) $\square$