CECL: Timely Loan Loss Provisioning and Bank Regulation*

Lucas Mahieux   Haresh Sapra   Gaoqing Zhang

Abstract

We investigate how provisioning models affect bank regulation. We study an accuracy vs. timeliness trade-off between an incurred loss model (IL) and a current expected credit loss model (CECL). Relative to IL, CECL improves efficiency by enabling timely intervention to curb inefficient *ex post* asset-substitution even though the imprecise information of CECL entails false alarms. However, from a real effects perspective, our analysis uncovers a potential cost of CECL: banks respond to timely intervention by originating riskier loans so that timely intervention induces timelier risk-taking. By appropriately tailoring regulatory capital to information about credit losses, the regulator can improve the efficiency of CECL. In particular, we show that regulatory capital under CECL would be looser when early estimates of credit losses are sufficiently precise and/or risk-shifting incentives are not too severe. From a policy perspective, our analysis suggests that better coordination between standard setters and bank regulators could enable the latter to relax capital requirements in order to spur lending.

Keywords: CECL; Expected Loss Model; Incurred Loss Model; Capital Requirements; Loan Loss Provisioning; Real Effects; Banking Regulation.

JEL codes: G21, G28, M41, M48.

---

1 Introduction

The recent adoption of the Current Expected Credit Loss standard (CECL) for the esti-mation of credit losses for loans and debt securities is arguably one of the most sweeping accounting changes to impact banks and financial institutions (Financial Accounting Stan-dards Board (FASB), 2019). Under CECL, banks would replace an "incurred loss" model with an "expected loss" model. A key difference between the two provisioning models is that under the incurred loss model, banks "delay recognition of credit losses until they have been incurred," whereas under the expected loss model, banks must recognize "the full amount of credit losses that are expected" as soon as loans are originated (FASB, 2016). While standard setters argue that the new standard would result in "more timely and relevant information," others have countered that it "could actually produce negative economic con-sequences" (Quaadman, 2019). Most notably, banks are concerned that the forecasts of future credit losses are often unreliable and false loss recognition may lower bank capital ratios and thereby "curtail credit availability, make credit losses worse during a recession and heighten volatility of bank earnings" (Maurer, 2020).[1] To shed light on this important debate, we develop an economic model to study the trade–offs in moving from an "incurred loss" model to an "expected loss" model.

We model a representative bank that is subject to shareholder–debtholder conflicts. The bank's shareholders have incentives to take excessive risk by either: 1) increasing the *ex*

---

[1]The COVID-19 crisis of 2020 is a case in point. There have been concerns that application of CECL during the crisis would severely erode banks' capital so that banks may curtail lending. In response, the CARES Act recently signed by President Trump on March 27, 2020 gives banks the option to delay imple-menting the new credit-loss standard until December 31, 2020, or until the end of the coronavirus national emergency, whichever comes first.

ante risk of the bank's loan portfolio by exerting less effort to screen borrowers, and/or 2) engaging in *ex post* asset substitution/risk-shifting to replace low-risk loans with high-risk ones. To discipline excessive risk-taking, a banking regulator imposes capital requirements. Importantly, the bank's level of capital depends on the provisioning model the bank uses to measure credit losses. Under an incurred loss model, the bank does not provision for credit losses until they are realized, whereas under the expected loss model, the bank relies on early forecasts of default risk to provision for credit losses upon loan origination. Provisioning for credit losses erodes the bank's level of regulatory capital, which, in turn, triggers regulatory intervention. The regulator then decides whether to take any regulatory action such as liquidating or restructuring the bank's loan portfolio.

We first study a benchmark setting in which the bank's *ex ante* choices of loan risk are kept fixed. We show that timely provisioning for loan losses under the expected loss model is always beneficial. Whenever the bank recognizes a loan loss in the interim, thereby eroding its regulatory capital, the regulator intervenes early and prevents the bank from engaging in inefficient asset substitution. More importantly, we show that the benefit of timely loss recognition always dominates the false alarm costs caused by the imprecise information inherent in the expected loss model. The reason is that a rational and benevolent regulator fully internalizes the false-alarm costs and acts optimally on the timely but imprecise information. In other words, our analysis shows that–as long as banks' *ex ante* risk choices are fixed–the accuracy-timeliness trade-off is one-sided as false alarm costs are not sufficient to overturn the benefits of timeliness offered by expected loss models.

However, once we allow for endogenous loan risk choices, early intervention is a double-edged sword. While timely intervention curbs *ex post* asset substitution, it induces the bank

to originate riskier loans *ex ante.* Originating riskier loans results in a surplus loss that may outweigh the benefits of timely loss recognition, thereby potentially making the expected loss model inferior to the incurred loss model. To understand this risk-aggravating effect of early intervention, note that under the incurred loss model, the bank prefers to defer taking excessive amounts of risk until it receives more precise information about the performance of its loan portfolio. However, under the expected loss model, the option value of waiting is constrained by regulatory intervention that preempts the bank from asset-substituting in the interim. Anticipating this, the bank responds by originating riskier loans *ex ante* so that timely intervention triggers even timelier risk-taking by the bank.

The overall efficiency of the expected loss model therefore depends on the trade-off between its *ex post* benefit of facilitating timely intervention and its *ex ante* cost of inducing excessive risk-taking. Interestingly, in incorporating the real effects of accounting measurements, more timely but less precise information under the expected loss model is no longer always socially desirable. When banks are heavily leveraged, the expected loss model becomes inferior. But when banks' leverage is not too high, the efficiency of the expected loss model hinges on its precision in estimating early loan losses. More precisely, the expected loss model is superior if interim signal generated by the model is sufficiently informative so that the false alarm costs of the expected loss model are not too high. Stated differently, the false alarm arguments made against expected loss models, i.e., the accuracy-timeliness trade-off, come to life *only* when we take into account the real effects of accounting measurements.

A policy implication of our findings is that changing the methodology for estimating credit losses for loans requires the banking regulator to simultaneously adjust capital adequacy ratios. We show that by appropriately tailoring the bank's regulatory capital ratio

to information about credit losses, the regulator can improve the efficiency of the expected loss model. In particular, our model suggests that relative to regulatory capital under an incurred loss model, regulatory capital under an expected loss model would be looser when the precision of estimating early credit losses is relatively high and/or when asset-substitution incentives are not too severe.

Our paper contributes to the banking literature on loan loss provisioning. Given the extensive size of this literature, we refer interested readers to two recent surveys by Beatty and Liao (2014) and Acharya and Ryan (2016). Most papers in this literature focus on empirically examining the effect of loan loss provisioning on banks' behavior. For instance, Beatty and Liao (2011) find that delays in loan loss recognition make banks' lending more procyclical.[2] Bushman and Williams (2012) document that timely recognition of expected future loan losses is associated with prudent risk-taking. Similarly, Bushman and Williams (2015) find that delayed expected loan loss recognition is positively associated with banks' vulnerability to stock market liquidity risk, downside tail risk of individual banks and codependency of downside tail risk among banks. Akins, Dou, and Ng (2019) report that timely loan loss recognition constrains corruption in bank lending and thus improves the quality of loans. We develop an analytical framework of the impact of loan loss provisioning on banks' risk-taking behavior and prudential regulation. In line with prior empirical evidence, we find that timely recognition of loan losses helps to curb inefficient *ex post* asset-substitution. However, our analysis also uncovers a potential cost of the timely loss recognition: it also induces banks to originate riskier loans *ex ante*. An empirical implication of our result is

---

[2]More recently, Laux and Rauter (2017) empirically investigate whether accounting measurements affect bank procyclicality.

that in examining the impact of loan loss provisioning, researchers should classify banks' risk decisions by the time they are undertaken (e.g., screening effort at loan origination versus asset substitution after learning bad news).

Our paper is also related to the burgeoning theoretical literature that examines the role of accounting measurements and disclosure in affecting financial stability and prudential regulation (see Goldstein and Sapra (2014) for a recent survey). Allen and Carletti (2008), Plantin, Sapra, and Shin (2008), Burkhart and Strausz (2009), and Mahieux (2019) examine the impact of mark-to-market accounting on bank risk and financial stability. Corona, Nan, and Zhang (2019) examine the coordination role of stress-test disclosure in affecting bank risk-taking. Gao and Jiang (2018), Zhang (2019), and Liang and Zhang (2019) study the role of accounting measurements in stabilizing bank runs. Finally, our study is closely related to several studies that also examine the use of accounting measurements in the prudential regulation of banks. Heaton, Lucas, and McDonald (2010) examine the interaction between mark-to-market accounting and capital requirements in affecting the social cost of regulation. Corona, Nan, and Zhang (2014) examine the effect of accounting information quality on the efficiency of capital requirements and banks' risk-taking incentives, taking into account the competition among banks. Corona, Nan, and Zhang (2019) and Bertomeu, Mahieux, and Sapra (2020) study the joint determination of optimal capital requirement policy and accounting measurement rules in disciplining banks' risk-taking and stimulating bank lending. Lu, Sapra, and Subramanian (2019) study the optimal use of mark-to-market accounting in implementing capital requirements, in the presence of asymmetric information and agency conflicts. However, unlike our study, neither loan loss provisioning nor its impact on prudential regulation are examined in any of the aforementioned studies. We shed light

$t = 0$	$t = 1$	$t = 2$	$t = 3$

Regulator sets capital requirements. Bank raises deposits D, and chooses effort q to screen borrowers.

Bank originates its loans. Bank learns expected loss s_1 and reports it according to loan loss provisioning rule. Regulator chooses intervention. Bank makes asset-substituting decision r.

Default risk s is realized. Regulator chooses intervention.
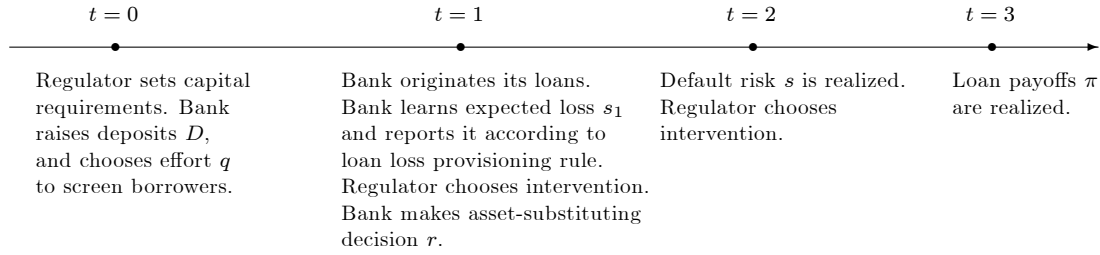
Loan payoffs π are realized.

Figure 1: Timeline of the model

on how regulatory capital should be tailored to the loan loss estimation model. More importantly, our analysis generates a new insight that relative to incurred loss models, optimal capital requirements under expected loss models may be looser when the precision of interim information are relatively high, and/or when the incentives to engage in inefficient asset substitution are not too severe. Our results therefore call for better coordination between banking regulators and accounting standard setters.

Section 2 describes our model. Section 3 provides an implementation of loan loss provisioning in the context of our model. Section 4 analyzes the model. Section 5 concludes. An Appendix contains the proofs of our results.

2 The Model

2.1 Timing of Events

We examine an environment that consists of a representative bank and a banking regulator. Figure 1 summarizes the timing of events.

At $t = 0$, the bank is endowed with an amount of exogenous equity E. The regulator chooses a capital ratio γ, defined as the equity-to-asset ratio for the bank. Given a fixed E, choosing the capital ratio γ is equivalent to choosing the size A of the bank's loan portfolio

or equivalently the bank size where $A \in \left[E, \bar{A}\right]$.[3] $\bar{A}$ denotes the maximum bank size and is chosen to be sufficiently large. For a bank of size $A > E$, the bank borrows $D = A - E$ from depositors. We assume that deposits are fully insured and we normalize the risk-free deposit rate to zero.

After raising deposits, the bank chooses a costly effort q to screen risky borrowers. The screening effort affects the return of the loan portfolio as follows. The outcome of each loan is binary: either the loan succeeds or it defaults. Absent any screening effort, i.e., $q = 0$, the bank always generates a high-risk loan that returns β with probability $\tau \in (0, 1)$ and 0 with probability $1 - \tau$. To improve the performance of the loan portfolio, the bank exerts screening effort q at a cost of $C(q)$ where the cost function $C(.)$ satisfies the standard properties: $C(0) = 0$, $C(1) = \infty$, $C'(0) = 0$, $C'(1) = \infty$, and $C'' > 0$. Conditional on a choice of $q > 0$, the bank receives the high-risk loan with probability $1 - q$ and the low-risk loan with probability q. The low-risk loan returns α with probability s and 0 with probability $1 - s$. The random variable s represents the (inverse) default risk of the low-risk loan and has a distribution $H(.)$ and a density $h(.)$ with full support on $[\tau, 1]$ so that the low-risk loan always has a lower default risk than the high-risk loan. This assumption guarantees that the bank's screening effort q reduces the default risk of the loan portfolio by decreasing the likelihood of high-risk loans in its loan portfolio. To reflect the risk-return trade-off, we assume that $\beta > \alpha > 1$ such that the bank demands a higher interest rate on the high-risk

---

[3] We consider the case that the regulator uses a simple leverage ratio to regulate capital. This is consistent with the newly proposed Basel III framework. In particular, Section V of Basel Committee (2010) provides a discussion of the use of leverage ratios in Basel III, that "the Committee agreed to introduce a simple, transparent, non-risk based leverage ratio that is calibrated to act as a credible supplementary measure to the risk based capital requirements." However, we do not explicitly study the exact implementation of the regulatory leverage requirement, which, in reality, could be achieved via minimum capital requirements, comprehensive capital analysis and review (CCAR), and other regulatory intervention actions.

loan.

At $t = 1$, after screening borrowers, the bank originates and invests A in its loans. The type of the loan is thereafter realized and privately observed by the bank. Conditional on originating a low-risk loan, the bank receives some early and potentially imprecise information s_1 regarding the default risk s. The random variable $\tilde{s}_1$ has a distribution $G(.)$ and a density $g(.)$ with full support on $[\tau, 1]$. We interpret s_1 as new information about a non-incurred loan value change that arrives at the intermediate date, and reflects the change in expectation of future loan losses. The arrival of such information affects the bank's assessment of the default likelihood and future loan losses. Formally, we model the bank's updated assessment of loan performance using the posterior distribution of s given s_1, denoted by $F(s|s_1)$ with a density $f(s|s_1)$. We assume that larger values of s_1 improve the bank's assessment of s in the sense of first-order stochastic dominance, i.e., $\frac{\partial F(s|s_1)}{\partial s_1} < 0$. Throughout our analysis, we use a general distribution for s and s_1. However, to illustrate our main results, following Corona, Nan, and Zhang (2019), we adopt the following linear-uniform structure:

$$\tilde{s} = w\tilde{s}_1 + (1 - w)\tilde{s}_2, \tag{1}$$

where $\tilde{s}_1$ and $\tilde{s}_2$ are i.i.d. random variables that are both uniformly distributed in the interval $[\tau, 1]$. The state $\tilde{s}$ is a weighted average of $\tilde{s}_1$ and $\tilde{s}_2$ where the weight on $\tilde{s}_1$ is $w \in [0, 1]$. The first component of $\tilde{s}_1$, is realized and observed by the bank at $t = 1$ while the second component $\tilde{s}_2$ is not realized until $t = 2$. In other words, observing s_1 at $t = 1$ reveals some partial information about s. The weight w thus captures the precision of s_1 in forecasting s. At $w = 0$, observing s_1 reveals no information regarding s, whereas at $w = 1$, observing s_1

perfectly reveals s.

An important goal of our model is to study the economic consequences of timely loan loss provisioning under an expected loss model such as CECL. We study two provisioning models. We refer to the model in which the bank does not provision for loan losses until the default risk s is realized as an "incurred loss model (IL)" and the model in which the bank additionally uses s_1 to provision for loan losses early as an "expected loss model (EL)." Relative to the incurred loss model, the expected loss model thus recognizes *a more timely but potentially imprecise estimate* of expected loan losses.

The timely recognition of loan losses facilitates the early regulatory intervention in the bank. In particular, under the incurred loss model, because the bank delays the recognition of loan losses, its capital ratio stays the same as $t = 0$, thus providing no basis for the regulator to intervene. In contrast, under the expected loss model, timely recognition of loan losses at origination leads to a write-down of the bank's loan, which, in turn, lowers its capital ratio. The erosion of the capital ratio, in turn, prompts regulatory intervention.[4] In practice, a regulator has broad discretion in terms of what actions to take ranging from being passive, thereby allowing the bank to continue its operations to a reorganization, partial asset sale, a reduction in the scope of the bank or even liquidation. For simplicity, we focus on two possible actions: continuation or liquidation.[5] We assume the regulator learns the type of the

---

[4]In Section 3, we provide a detailed account on the implementation of loan loss provisioning and its impact on a bank's balance sheet and its capital ratio. In particular, we illustrate how provisioning for loan losses triggers regulatory intervention under the incurred loss vs. the expected loss models.

[5]The structure of *ex-post* intervention actions (continuation vs. liquidation) that our model adopts follows directly from the continue-stop modeling structure ("C" vs. "S") developed in the seminal work of Dewatripont and Tirole (1994). For convenience, we refer to the action that reduces the risk of the bank's loans as "liquidation." Note, however, that this action is not only limited to liquidating the bank's loans, but may also be broadly interpreted as any action that makes the terminal cash flows of the loans less volatile and more deterministic, for instance, restructuring the bank's loan portfolio, etc.

loan at the intervention stage. Given the information s_1, the regulator optimally liquidates the bank's loan if the expected payoffs from liquidation exceed the expected payoffs from continuation. If a unit of low-risk loan is liquidated, the regulator recovers the original investment of 1 whereas if a unit of high-risk loan is liquidated, the regulator recovers L_β. In case of continuation at $t = 1$, based on its updated assessment of the loan performance $F(s|s_1)$, the bank may increase the risk of the loan portfolio by engaging in asset substitution $r \in \{0, 1\}$. The variable $r = 0$ implies no asset substitution so that the bank does not change its original loan portfolio whereas the variable $r = 1$ implies that the bank changes its original loan portfolio by substituting the low-risk loans in the loan portfolio with high-risk loans. That is, conditional on $r = 1$, the bank always receives the high-risk loan.

At $t = 2$, conditional on continuation of the bank at $t = 1$, the default risk s is realized and observed by the bank. Since the default risk is realized, the bank reports s, regardless of whether the bank follows the incurred loss model or the expected loss model. The recognition of loan losses reduces the bank's capital ratio, in which case the regulator decides, again, whether to intervene.

At $t = 3$, the terminal payoffs of the loans are realized. The per-unit payoff π of the loan portfolio is as follows. If the loan is low-risk and is liquidated, $\pi = 1$ but if the loan is continued, $\pi = \alpha$ with probability s and $\pi = 0$ with probability $1 - s$. If the loan is high-risk and is liquidated, $\pi = L_\beta$ whereas in case of continuation, $\pi = \beta$ with probability τ and $\pi = 0$ with probability $1 - \tau$. The regulator compensates depositors if the bank fails, i.e., $\pi < D$, with a lump sum payment which we assume is financed via a frictionless *ex ante* tax.

2.2 Assumptions

To create a demand for prudential regulation, we impose the following assumptions (A1-A4) throughout our entire analysis:

Assumption 1: The regulator cannot commit to a particular intervention action.

This assumption is consistent with the practice of prudential regulation (see e.g., Dewatripont and Tirole, 1994). For instance, in the United States, the Federal Reserve only sets capital requirement ratios (e.g., to be well-capitalized, a bank holding company must have a Tier-1 capital ratio of 6%) but never firmly specifies any intervening action. Such lack of commitment by banking regulators has been examined extensively in the literature (e.g., Bagehot, 1873; Mailath and Mester 1994; Freixas, 1999).

Assumption 2: Liquidating high-risk loans results in a lower payoff than continuation in expectation, i.e.,

$$L_\beta < \tau\beta. \tag{2}$$

Two points regarding Assumption 2 are noteworthy. First, it captures realistic features of markets for risky loans such as sub-prime loans: these markets are illiquid and consist of buyers who are typically the second-best users so that regulators often recover very little residual value. (Acharya and Yorulmazer, 2007)[6]. Therefore, unlike liquidating prime loans when regulators typically recover the original investment, liquidating risky subprime loans entails significant losses. Second, it implies that in our model, once the bank has already originated the high-loan loan, the regulator cannot commit to intervening in the bank because

---

[6]Assumption 2 can be relaxed to incorporate random liquidation values that may be higher or lower than the continuation value. What is important for our model is that the *expected liquidation value* of the high-risk loan is lower than its continuation value so that liquidating subprime loans entails losses in *expectation.*

liquidating the bank's high-risk loan entails large expected losses. Such *ex-post* inefficiency in the regulatory intervention, in turn, calls for the optimal design of the *ex-ante* capital requirement policy to deter the bank from originating high-risk loans in the first place. Stated differently, under Assumption 2, the capital requirement policy becomes an essential tool for the regulator in disciplining banks' excessive risk-taking incentives.

Assumption 3: A high-risk loan is value-destroying whereas a low-risk loan is value-creating.

This assumption ensures that the regulator has an incentive to discipline the bank from investing in high-risk loans. It turns out that a sufficient condition for Assumption 3 is

$$\tau\beta < 1. \tag{3}$$

Note that because the low-risk loan generates a liquidation payoff of 1, it generates at least zero NPV even in case of liquidation. In addition, (3) ensures that the high-risk loan always generates negative NPV.[7]

Assumption 4: The bank always prefers to invest in the high-risk loan if it lends the maximal extent, i.e., $A = \bar{A}$.

This assumption rules out the degenerate case in which the bank can achieve the first-best by choosing the highest leverage and lending the maximal extent, thus making the capital requirement regulation undesirable. When $\bar{A}$ is sufficiently large, this assumption is reduced

---

[7]Assuming that both loans generate positive NPV does not affect our analysis qualitatively. Given the constraint on the bank's size, $A \leq \bar{A}$, the regulatory problem would then reduce into one in which the regulator has an incentive to induce the bank to invest in the loan with a higher NPV.

into:

$$\tau \left(\beta - 1 \right) > \left(\int_{\frac{1}{\alpha}}^{1} s f \left(s | s_1 \right) ds \right) \left(\alpha - 1 \right), \tag{4}$$

for any $s_1 \in [\tau, 1]$.

3 Provisioning for loan losses and capital ratio

We now provide an implementation of loan loss provisioning accounting standards under CECL in the context of our model. In particular, we show how provisioning for loan losses affects the bank's balance sheet and capital ratio.[8] For simplicity, we assume that the events at $t = 3$ occur immediately after those at $t = 2$ to avoid the recognition of the accrual of interest revenue between $t = 2$ and $t = 3$. Stated differently, we assume that the length of time between the date at which the uncertainty regarding the default risk is resolved ($t = 2$) and the date at which the loan's terminal cash flows are realized ($t = 3$) is negligible relative to the entire duration of the loan (i.e., from origination at $t = 1$ to maturity at $t = 3$). We use the notations E_t and A_t to denote the carrying values of the bank's equity and assets at date t, respectively.

At $t = 0$, the value of the bank's assets $A_0 = E + D = A$ and the value of the bank's

---

[8] We are grateful to Mary Barth and Alexander Nezlobin who encouraged us to better connect our model to the practical implementation of an expected loss model such as CECL.

equities $E_0 = E$. The regulatory capital requirement thus requires that[9]:

$$\frac{E_0}{A_0} = \frac{E}{A} = \gamma. \tag{5}$$

At $t = 1$, after originating the loans, the bank considers whether to provision for loan losses. Consider first the case that the bank originates a low-risk loan. Under the incurred loss model, because the bank delays the recognition of loan losses, its asset and equity values are unchanged, i.e., $A_1 = A$ and $E_1 = E$. The bank continues to satisfy the regulatory capital requirement,

$$\frac{E_1}{A_1} = \frac{E}{A} = \gamma. \tag{6}$$

As a result, the regulator cannot intervene at $t = 1$ under the incurred loss model.

The recognition of loan losses under the expected loss model, however, differs substantially from that under the incurred loss model because the expected loss model "forces banks to recognize expected future losses immediately" at origination.[10] In particular, FASB Accounting Standards Update No. 2016-13 states that, upon initial measurement, "the entity shall discount expected cash flows at the financial asset's effective interest rate," and "the allowance for credit losses shall reflect the difference between the amortized cost basis and the present value of the expected cash flows." (FASB, 2016, paragraph 326-20-30-4). Under

---

[9]Bank regulators may also use other types of regulation based on the bank's balance sheet such as some risk-weighted measure of assets to regulate capital. By Basel II and III, assets are partitioned into different groups based on their risk and these different groups are assigned different weights. Adding an additional risk-weighted capital constraint in our model, however, will not alter any of our results for two reasons. First, at $t = 0$, the bank always satisfies the risk-weighted capital requirement because, prior to originating its loans, the bank's assets consists of cash that receives a risk weight of 0. Second, as shown in our later analysis, with only the leverage requirement, the regulator already implements the *ex post* optimal intervention policy at $t = 1$ and $t = 2$. Adding another regulatory constraint hence alters neither the regulator's intervention decisions nor the bank's decisions.

[10]See https://bpi.com/cecl-regulatory-capital-proposal-leaves-many-important-questions-unanswered/.

the expected loss model, the bank therefore accounts for the expected change s_1 in the default risk and measures the value of the loan by discounting the expected future cash flows from the loan $(A\alpha E\left[s|s_1\right])$ using the loan's effective interest rate $(\alpha - 1)$:

$$A_1 = \frac{A\alpha E\left[s|s_1\right]}{1+\alpha-1} = AE\left[s|s_1\right]. \tag{7}$$

Comparing A_1 to A_0 implies that the balance of the allowance for loan losses, $A_0 - A_1 = A\left(1 - E\left[s|s_1\right]\right)$. Provisioning for loan losses, in turn, reduces the bank's equity to $E_1 = E - A\left(1 - E\left[s|s_1\right]\right)$ so that the bank's capital ratio changes to

$$\frac{E_1}{A_1} = \frac{E - A\left(1 - E\left[s|s_1\right]\right)}{A - A\left(1 - E\left[s|s_1\right]\right)} < \frac{E}{A} = \gamma. \tag{8}$$

That is, under the expected loss model, reporting s_1 at $t = 1$ violates the regulatory capital requirement that, in turn, triggers regulatory intervention.

Consider next the case that the bank originates a high-risk loan. Since the default risk τ is realized at $t = 1$, under both the incurred loss model and the expected loss model, the bank measures the value of the loan by discounting the expected future cash flows from the loan $(A\beta\tau)$ to their present value using the loan's effective interest rate $(\beta - 1)$:

$$A_1 = \frac{A\beta\tau}{1+\beta-1} = A\tau. \tag{9}$$

Comparing A_1, the carrying value of the loan at $t = 1$ to A_0, the carrying value of the loan at $t = 0$ gives the amount of impairment on the loan, $A_0 - A_1 = A\left(1 - \tau\right)$. After provisioning for loan losses, the bank's equity becomes $E_1 = E - A\left(1 - \tau\right)$ so that the bank's capital

ratio becomes

$$\frac{E_1}{A_1} = \frac{E - A(1 - \tau)}{A - A(1 - \tau)} < \frac{E}{A} = \gamma. \tag{10}$$

Under both incurred loss and expected loss, the bank who invests in a high-risk loan always violates the regulatory capital requirement at $t = 1$. However, since liquidating the high-risk loan entails losses in expectation, the regulator will choose to continue the bank's high-risk loan in equilibrium so that the violation of capital requirements does not matter.

Finally, at $t = 2$, if the bank originates a low-risk loan, the recognition of loan losses is similar to that in $t = 1$, except that the bank needs to recognize the expected interest revenue $A\alpha(s - 1)$ accrued between $t = 1$ and $t = 2$. In particular, under both the incurred and the expected loss models, the bank recognizes the accrued interest revenue and measures the value of the loan at the total expected cash flows from the loan:

$$A_2 = As + As(\alpha - 1) = A\alpha s. \tag{11}$$

Comparing A_2 to the carrying value of the loan at $t = 0$ gives the adjusted balance of the allowance for loan losses: $A_0 - A_2 = A(1 - \alpha s)$. Adjusting the loan losses, in turn, changes the value of the bank's equity, i.e., $E_2 = E - A(1 - \alpha s)$. The bank's capital ratio thus changes to

$$\frac{E_2}{A_2} = \frac{E - A(1 - \alpha s)}{A - A(1 - \alpha s)}. \tag{12}$$

Note that $\frac{E_2}{A_2} < \frac{E}{A} = \gamma$ if and only if $\alpha s < 1$. Therefore, the bank with a low-risk loan violates the capital requirement at $t = 2$ if and only if $\alpha s < 1$.

If the bank originates a high-risk loan, under both the incurred and the expected loss

models, the bank measures the value of the loan at the total expected cash flows from the loan, including the accrued expected interest revenue $A\tau(\beta - 1)$:

$$A_2 = A\tau + A\tau(\beta - 1) = A\beta\tau. \tag{13}$$

Comparing A_2 to the carrying value of the loan at $t = 0$ gives the adjusted amount of the allowance for loan losses: $A_0 - A_2 = A(1 - \beta\tau)$. After recognizing the loan losses, the bank's equity becomes $E_2 = E - A(1 - \beta\tau)$ so that the bank's capital ratio changes to

$$\frac{E_2}{A_2} = \frac{E - A(1 - \beta\tau)}{A - A(1 - \beta\tau)} < \frac{E}{A} = \gamma. \tag{14}$$

The inequality is due to $\beta\tau < 1$. Therefore, the bank that originates a high-risk loan always violates the capital requirement at $t = 2$.

In sum, at $t = 1$, the bank with a low-risk loan violates the regulatory capital requirement under the expected loss model but satisfies it under incurred loss. As a result, under incurred loss, the regulator cannot intervene at $t = 1$ whereas, under expected loss, the bank's violation of the capital requirements allows the regulator to implement the *ex post* optimal intervention policy, as assumed in our model setup. At $t = 2$, under both loan loss models, the bank satisfies the regulatory capital requirement if and only if the bank invests in a low-risk loan and $\alpha s \geq 1$. Note that these triggers of capital requirements violations at $t = 2$, again, allow the regulator to implement the *ex post* optimal intervention policy to the extent that whenever the bank meets the capital requirement, the regulator also prefers to continue the bank *ex post*.

4 Analysis

4.1 Exogenous loan risk

As a benchmark, we first solve our model for a given loan portfolio, i.e., treating the bank's *ex ante* risk q as exogenous.

4.1.1 Incurred loss model with exogenous risk

We analyze the incurred loss model in which the bank does not recognize s_1 so that the regulator can only intervene at $t = 2$ after the default risk s is realized. We solve the model using backward induction. At $t = 2$, given default risk s, the regulator decides whether to liquidate the bank by comparing the total expected payoffs from liquidation with the total payoffs from continuation. If the bank's loan turns out to be high-risk, it generates L_β upon liquidation and $\tau\beta > L_\beta$ upon continuation. Therefore, the regulator never liquidates the high-risk loan. On the other hand, if the loan turns out to be low-risk, it generates 1 upon liquidation and αs upon continuation. Therefore, the regulator liquidates the low-risk loan if and only if

$$s < \frac{1}{\alpha}. \tag{15}$$

Next, we solve for the bank's asset-substituting decision r at $t = 1$ conditional on the early information s_1. Denote the bank's shareholders' (hereafter the bank's) expected payoff by $U(r|s_1)$. If the bank has already received a high-risk loan, the asset-substitution decision is moot. If the bank receives a low-risk loan and chooses not to engage in asset substitution

$(r = 0)$, its payoff is:

$$U(0|s_1) = \left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1)\, ds \right] (A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] E. \tag{16}$$

Note that the bank's expected payoff at $t = 1$ depends on the regulator's liquidation decision at $t = 2$. If $s \geq 1/\alpha$ so that the regulator does not liquidate the bank, the bank receives a net payoff of $A\alpha - (A - E)$ after repaying depositors with probability s, and receives 0 with probability $1 - s$. But if $s < 1/\alpha$, the regulator liquidates the bank so that the bank receives the certain liquidation payoff $A - (A - E) = E$ after repaying depositors.

If the bank engages in asset substitution so that $r = 1$, its payoff is

$$U(1|s_1) = \tau(A\beta - (A - E)). \tag{17}$$

Conditional on obtaining a high-risk loan, the bank receives $A\beta - (A - E)$ after repaying depositors with probability τ, and receives 0 with probability $1 - \tau$. Note that since the bank's payoff from a high-risk loan $U(1|s_1)$ does not depend on s_1, to economize on notations, we hereafter omit s_1 in $U(1|s_1)$ and denote it by $U(1)$. The bank chooses asset substitution if and only if

$$U(1) > U(0|s_1), \tag{18}$$

which reduces into

$$\tau(A\beta - (A - E)) > \left(\left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1)\, ds \right] (A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] E \right). \tag{19}$$

20

The left hand side of (19) represents the bank's payoff from investing in a high-risk loan and does not depend on s_1. As s_1 increases, the posterior distribution of s shifts to the right so that the right hand side of (19) which represents the bank's payoff from investing in a low-risk loan strictly increases in s_1. As a result, there exists a unique cutoff $\bar{s}_1(A)$ such that the bank engages in asset substitution if and only if its early information regarding the low-risk loan's performance is below the cutoff, i.e., $s_1 < \bar{s}_1(A)$. We next formally state the bank's equilibrium asset-substituting decision.

Proposition 1 *There exists a unique threshold $\bar{s}_1(A) \in [\tau, 1]$ such that the bank makes the asset-substituting decision $(r = 1)$ if and only if $s_1 < \bar{s}_1(A)$. The threshold $\bar{s}_1(A)$ is given by:*

$$
\begin{cases}
\bar{s}_1(A) = \tau & \text{if } A \in [E, A_{\min}]; \\
\\
\bar{s}_1(A) \in (0, 1) & \text{if } A \in (A_{\min}, A_{\max}); \\
\\
\bar{s}_1(A) = 1 & \text{if } A \in [A_{\max}, \bar{A}].
\end{cases}
$$

The cutoffs $\{A_{\min}, A_{\max}\}$, where $A_{\min} < A_{\max}$, are defined in the Appendix. For $A \in (A_{\min}, A_{\max})$, $\bar{s}_1(A)$ is the unique solution to:

$$
\tau = \left[\int_{\frac{1}{\alpha}}^{1} s f\left(s | \bar{s}_1(A)\right) ds \right] \frac{A\alpha - (A - E)}{A\beta - (A - E)} + \Pr\left[s < \frac{1}{\alpha} | \bar{s}_1(A) \right] \frac{E}{A\beta - (A - E)}. \tag{20}
$$

Ceteris paribus, the threshold $\bar{s}_1(\cdot)$ increases in the bank size A and in the face value of the high-risk loan β but decreases in the face value of the low-risk loan α for $\bar{s}_1(A) \in (0, 1)$.

Proposition 1 is intuitive. It states that when the bank expects the performance of its loan portfolio to deteriorate, its incentives to engage in asset substitution increase. Furthermore,
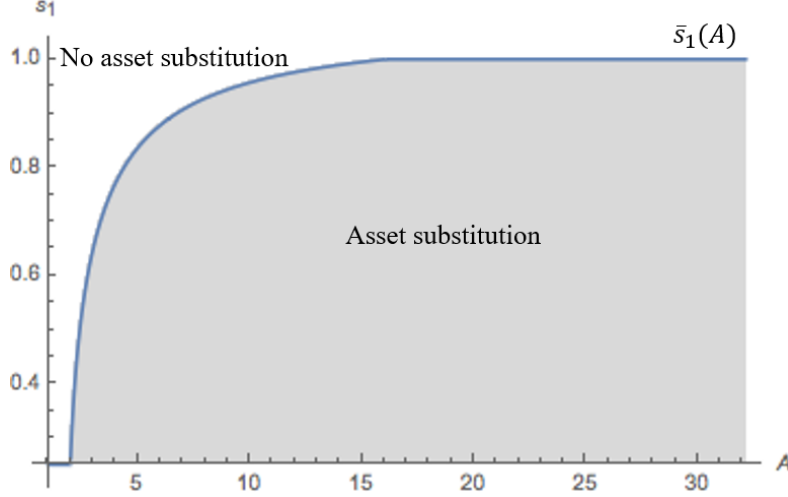
Figure 2: The effect of leverage on the asset-substituting threshold for $A \in \left[E, \bar{A}\right]$. The following parameter values are used in this plot: $\alpha = 1.5$, $\beta = 2.5$, $E = 1$, $\tau = 0.25$ and $w = 0.45$.

such incentives become sharper when either the relative payoff of the high-risk loan to the low-risk loan β/α increases and/or the bank's leverage A increases. The latter result suggests a beneficial role for capital requirements in curbing asset substitution: a higher regulatory capital ratio γ (equivalently, a lower A) weakens the bank's asset-substituting incentives in the interim. Figure 2 illustrates the effect of A on the asset-substitution decision. When the leverage of the bank is extremely low (i.e., whenever the bank faces a tight capital requirement), the bank never engages in asset substitution, whereas when the leverage is extremely high (whenever the bank faces a loose capital requirement), the bank always chooses to asset-substitute. For intermediate values of leverage, the bank's assessment of its loan performance matters and the bank engages in asset substitution if and only if such assessment deteriorates.

4.1.2 Expected loss model with exogenous risk

We now analyze the expected loss model in which the bank reports s_1 so the regulator can either intervene early at $t = 1$ based on s_1 or late at $t = 2$ based on s. We again solve the model using backward induction. Note that the regulator's liquidation decision at $t = 2$ is the same as that in the incurred loss model because under both models, the regulator receives the full information s about the default risk. Facing the same liquidation policy at $t = 2$, the bank's asset-substituting decision also stays the same across both models. Nonetheless, there exists a key difference between the two loan loss models: under the expected loss model, timely recognition of loan losses s_1 allows the regulator to intervene early at $t = 1$.

More specifically, conditional on a high-risk loan, the regulator intervenes but chooses to not liquidate, as explained previously. Conditional on a low-risk loan, two key factors come into play when the regulator decides whether to liquidate: 1) the information s_1 and 2) the bank's future asset-substituting decision r. While the regulator receives a constant payoff of 1 from liquidation, her expected payoff from continuation depends on both s_1 and r. If $s_1 \geq \bar{s}_1(A)$, the regulator rationally anticipates that the bank will keep the low-risk loan whose expected surplus is

$$\int_{\frac{1}{\alpha}}^{1} \alpha s f\left(s | s_1\right) ds + \Pr\left[s < \frac{1}{\alpha} | s_1\right] > 1. \tag{21}$$

That is, absent the asset-substitution problem, the regulator should never liquidate the bank early based on the timely but imperfect information s_1, because such information may entail false alarms (i.e., liquidating a bank whose expected payoff of continuation exceeds that of liquidation, $\alpha s > 1$). The regulator can always avoid the false alarm cost by postponing

23

the decision to a later date when better information arrives (i.e., the default risk s is fully realized at $t = 2$).

However, given the bank's asset substitution incentives when $s_1 < \bar{s}_1(A)$, early intervention becomes desirable to curb such behavior. To elaborate, consider the following trade-off the regulator faces in intervening early in the bank's operations. On the one hand, as discussed previously, there are still false alarm costs from early intervention as it relies on imprecise information s_1 about future default risk s. The regulator may liquidate a bank that is financially sound $(s > 1/\alpha)$ but nevertheless receives some bad interim information $(s_1 < \bar{s}_1(A))$, resulting in false alarm costs of

$$\int_{\tau}^{\bar{s}_1} (\alpha s - 1) \, g(s_1) \, ds_1. \tag{22}$$

On the other hand, however, there is also a benefit of early intervention stemming from curbing asset substitution. To see this, note that when the loan performance deteriorates, i.e., $s_1 < \bar{s}_1(A)$, the regulator anticipates that the bank will switch to the high-risk loan, which generates an expected surplus of $\tau\beta$. The asset substitution, in turn, results in an expected surplus loss relative to the expected surplus that the low-risk loan could have generated had the loan been continued which is given by:

$$\int_{\tau}^{\bar{s}_1} (\alpha s - \tau\beta) \, g(s_1) \, ds_1. \tag{23}$$

In equilibrium, the regulator decides whether to intervene early by comparing the false alarm cost with the expected surplus loss from asset substitution. The regulator intervenes if and

only if the asset-substitution surplus loss outweighs the false alarm cost, which always holds because

$$\int_{\tau}^{\bar{s}_1} (\alpha s - \tau \beta) g(s_1) ds_1 - \int_{\tau}^{\bar{s}_1} (\alpha s - 1) g(s_1) ds_1$$
$$= \int_{\tau}^{\bar{s}_1} (1 - \tau \beta) g(s_1) ds_1 \geq 0. \tag{24}$$

The last inequality follows from $1 > \tau \beta$ (Assumption 3). That is, by curbing asset substitution, early intervention always generates *net* surplus gains despite the false alarm cost. The term in (24) thus represents the net social benefit of timely intervention. Stated differently, a rational and benevolent regulator fully internalizes the false alarm costs from acting on early but imprecise information and chooses an action that results in the highest surplus *in expectation.* We formally state the regulator's decision at $t = 1$ in the following proposition.

Proposition 2 *Under the expected loss model, conditional on the early information s_1, the regulator intervenes and liquidates the bank at $t = 1$ if and only if the bank has a low-risk loan and $s_1 < \bar{s}_1(A)$.*

4.1.3 Surplus comparison with exogenous risk

With the equilibrium characterized for a given bank size A, we compare the surplus between under the incurred loss model and the expected loss model, holding the bank's *ex ante* risk

25

choice q fixed. The surplus under incurred loss is

$$
\begin{aligned}
W_{IL}(A) &= (1-q)A\tau\beta + q\Big(A \int_{\tau}^{\bar{s}_1(A)} \tau\beta g(s_1)ds_1 \\
&\quad + A \int_{\bar{s}_1(A)}^{1} \Big(\Big[\int_{\frac{1}{\alpha}}^{1} sf\left(s|s_1\right)ds \Big] \alpha + \Pr\Big[s < \frac{1}{\alpha}|s_1 \Big] \Big) g(s_1)ds_1 \Big) - A. \quad (25)
\end{aligned}
$$

With probability $1-q$, the bank gets a high-risk loan and the expected surplus is $\tau\beta$, whereas with probability q, the bank gets a low-risk loan. The bank keeps the low-risk loan if $s_1 \geq \bar{s}_1(A)$ and switches to a high-risk one if $s_1 < \bar{s}_1(A)$.

Under the expected loss model, the surplus is

$$
\begin{aligned}
W_{EL}(A) &= (1-q)A\tau\beta + q\Big(A \int_{\tau}^{\bar{s}_1(A)} g(s_1)ds_1 \\
&\quad + A \int_{\bar{s}_1(A)}^{1} \Big(\Big[\int_{\frac{1}{\alpha}}^{1} sf\left(s|s_1\right)ds \Big] \alpha + \Pr\Big[s < \frac{1}{\alpha}|s_1 \Big] \Big) g(s_1)ds_1 \Big) - A. \quad (26)
\end{aligned}
$$

The only difference between the incurred and the expected loss models is that, if the bank gets a low-risk loan which occurs with probability q and $s_1 < \bar{s}_1(A)$, the bank is liquidated under the expected loss model whereas the bank is allowed to continue and asset-substitute under the incurred loss model. Recognizing this difference, we obtain:[11]

$$
W_{EL}(A) - W_{IL}(A) = qA \left[\int_{\tau}^{\bar{s}_1(A)} (1 - \tau\beta)\, g(s_1)ds_1 \right] \geq 0. \quad (27)
$$

Since the expected loss model dominates for all ranges of asset sizes A, it also dominates when the regulator sets the capital requirement policy optimally. We formally state this

---

[11]The inequality is strict if $\bar{s}_1(A) > \tau$ (i.e., $A \in (A_{\min}, \bar{A}]$ from Proposition 1). Note that the surplus gain from expected loss in (27) coincides with the net benefit of timely intervention in (24).
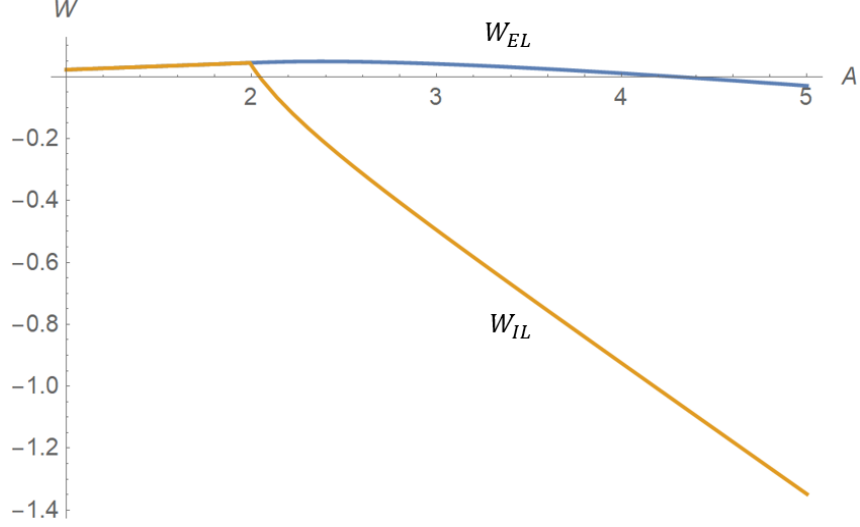
Figure 3: Surplus comparison with exogenous risk for $A \in [E, \bar{A}]$. The following parameter values are used in this plot: $\alpha = 1.5$, $\beta = 2.5$, $E = 1$, $\tau = 0.25$, $w = 0.45$, and $q = 0.9$.

result in the following proposition and provide a graphical illustration in Figure 3.

Proposition 3 *Fixing the bank's ex ante risk q of the loan portfolio, the expected loss model always dominates the incurred loss model.*

Proposition 3 implies that the early information revealed by the expected loan loss model generates an expected benefit because it allows regulators to intervene in a more timely manner in banks' operations to curb inefficient asset substitution. Interestingly, such benefit cannot be overturned by the false-alarm costs stemming from the imprecise information of the expected loss model, thereby making accuracy vs. timeliness trade–off one-sided. The reason is that, the regulator rationally takes into account the (im)precision of the early information and therefore internalizes such false alarm costs in determining whether to intervene. The preceding result supports claims made by proponents of expected loss models who have argued that by providing more timely information about the performance of banks' loans, expected loss models would prompt earlier corrective action in bad times. Proposition 3

confirms those views and shows that expected loss models dominate incurred loss models *as long as the ex ante risk of the banks' loan portfolios is kept fixed.* However, banks are not passive technologies–rather banks' insiders could respond to the regulator's intervention strategy by changing their lending behavior. We next investigate the impact of provisioning models on the *ex ante* risk of the loans that banks originate.

4.2 Endogenous loan risk

We now analyze the complete model in which the bank can choose the riskiness q of the loans at the origination stage.

4.2.1 Incurred loss model with endogenous risk

We start with the incurred loss model. For a given bank size A, the bank chooses risk q that solves

$$q_{IL}^* \in \arg \max_{q \in [0,1]} U(q) = \underbrace{\int_{\tau}^{\bar{s}_1(A)} U(1) g(s_1) \, ds_1}_{\text{expected payoff given asset substitution}} \tag{28}$$

$$+ \underbrace{\int_{\bar{s}_1(A)}^{1} [q U(0|s_1) + (1-q) U(1)] g(s_1) \, ds_1}_{\text{expected payoff given no asset substitution}} - C(q).$$

Recall that if $s_1 < \bar{s}_1(A)$, the bank's payoff is independent of its initial risk choice because the bank will engage in asset substitution and change the loan portfolio into a high-risk one. Thus, the *ex ante* risk choice only matters when the bank does not asset substitute. The higher the likelihood of interim asset substitution, the lower the bank's incentives to engage in costly *ex ante* screening, i.e., q_{IL}^* decreases in $\bar{s}_1(A)$. As the bank's leverage becomes

very large, i.e., $A \in [A_{\max}, \bar{A}]$ so that $\bar{s}_1(A) = 1$, the bank always engages in interim asset substitution making the *ex ante* risk decision moot. In this case, because screening is costly, the bank chooses not to screen the borrowers *ex ante* so that $q_{IL}^* = 0$.

The first order condition of the preceding equation with respect to q yields:

$$\int_{\bar{s}_1(A)}^1 [U(0|s_1) - U(1)] g(s_1) ds_1 = C'(q_{IL}^*). \tag{29}$$

The right hand side of equation (29) captures the marginal cost of screening borrowers whereas the left hand side captures the marginal benefit of screening stemming from reducing the future default risk. To see the latter effect, note that from equation (19),

$$U(0|s_1) - U(1) = \left(\left[\int_{\frac{1}{\alpha}}^1 s f(s|s_1) ds \right] (A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha} \Big| s_1 \right] E \right) - \tau(A\beta - (A-E)), \tag{30}$$

which is the incremental payoff from investing in the low-risk loan vs. the high-risk loan. It is positive because, by the incentive-compatibility constraint (19), the bank strictly prefers the low-risk loan for $s_1 \geq \bar{s}_1$.

We formally state the bank's *ex ante* risk choice in the following proposition.

Proposition 4 *Under the incurred loss model, the bank chooses risk q_{IL}^* such that, for bank size $A \in [E, A_{\max})$, $q_{IL}^* \in (0,1)$ where q_{IL}^* solves equation (29), whereas for $A \in [A_{\max}, \bar{A}]$, $q_{IL}^* = 0$. Ceteris paribus, the risk choice q_{IL}^* decreases in bank size A, increases in the face value of the low-risk loan α, and decreases in the face value of the high-risk loan β.*

4.2.2 Expected loss model with endogenous risk

Under the expected loss model, the bank chooses risk q that solves

$$q^*_{EL} \in \arg \max_{q \in [0,1]} U(q) = \int_\tau^{\bar{s}_1(A)} [qE + (1-q)U(1)] g(s_1) ds_1 \tag{31}$$
$$+ \int_{\bar{s}_1(A)}^1 [qU(0|s_1) + (1-q)U(1)] g(s_1) ds_1 - C(q).$$

Note that when $s_1 < \bar{s}_1(A)$, the bank's expected payoff under the expected loss model differs from that under the incurred loss model in (28). Under the expected loss model, if the regulator expects the performance of the bank's portfolio to deteriorate, the regulator intervenes and disciplines the bank with the low-risk loan by preventing it from engaging in asset substitution. Using this equilibrium property, we obtain that for $s_1 < \bar{s}_1(A)$, the bank's expected payoff is $qE + (1-q)U(1)$.

Taking the first order condition with respect to q yields:

$$\int_\tau^{\bar{s}_1(A)} [E - U(1)] g(s_1) ds_1 + \int_{\bar{s}_1(A)}^1 [U(0|s_1) - U(1)] g(s_1) ds_1 = C'(q^*_{EL}). \tag{32}$$

To compare the risk choices across the provisioning models, we plug the first order condition (29) on q^*_{IL} into the first order condition (32) to obtain:

$$C'(q^*_{EL}) - C'(q^*_{IL}) = \int_\tau^{\bar{s}_1(A)} [E - U(1)] g(s_1) ds_1 \le \int_\tau^{\bar{s}_1(A)} [U(0|s_1) - U(1)] g(s_1) ds_1 \le 0, \tag{33}$$

where all the inequalities are strict if $\bar{s}_1 > \tau$. The first inequality holds because, due to its limited liability, the bank strictly prefers the regulator to continue the low-risk loan than

liquidate it early.[12] The second inequality holds because, by the incentive-compatibility constraint (19), the bank strictly prefers the high-risk loan for $s_1 \leq \bar{s}_1$. Since $C'' > 0$, $q_{EL}^* \leq q_{IL}^*$.

A key message from equation (33) is that–when real effects are taken into account–early intervention is a double-edged sword: while timely intervention curbs asset substitution *ex post*, it might induce timelier risk-taking *ex ante*. To understand the latter risk-aggravating effect, note that under the incurred loss model, the bank prefers to defer taking excessive amounts of risk until it receives more precise information on the performance of its loan portfolio (i.e., upon learning $s_1 \leq \bar{s}_1$). However, under the expected loss model, the option value of waiting for more precise information diminishes due to regulatory intervention that preempts asset substitution in the region of $s_1 \in [\tau, \bar{s}_1]$. Anticipating this, the bank shifts the timing of its risk-taking earlier by originating riskier loans *ex ante* under the expected loss model.

Finally, when A increases and becomes sufficiently close to $A_{\max}$, then $\lim_{A \to A_{\max}} \bar{s}_1(A) = 1$ so that the left hand side of the first order condition (32) becomes

$$\int_\tau^1 [E - U(1)] g(s_1) ds_1 < 0. \tag{34}$$

The inequality holds because, by (33), for $s_1 \leq \bar{s}_1(A_{\max}) = 1$, $E - U(1) \leq 0$ so that at $A = A_{\max}$, $q_{EL}^* = 0$. In addition, when A is sufficiently small and close to $A_{\min}$, $\lim_{A \to A_{\min}} \bar{s}_1(A) = \tau$, and by (33), $q_{EL}^* = q_{IL}^* > 0$. An application of the intermediate value theorem thus

---

[12]Mathematically, $U(0|s_1) - E = \left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1) ds \right] (A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|s_1\right] E - E = \int_{\frac{1}{\alpha}}^1 [s(A\alpha - (A - E)) - E] f(s|s_1) ds > \frac{1}{\alpha}(A\alpha - (A - E)) - E = (A - E)\left(1 - \frac{1}{\alpha}\right) > 0.$

31

suggests that, there exists some cutoff $A_e \in (A_{\min}, A_{\max})$ such that for $A \geq A_e$, we have $q^*_{EL} = 0$. The cutoff A_e is defined by

$$\int_\tau^{\bar{s}_1(A_e)} [E - U(1)]\, g(s_1)\, ds_1 + \int_{\bar{s}_1(A_e)}^1 [U(0|s_1) - U(1)]\, g(s_1)\, ds_1 = 0, \qquad (35)$$

such that the left hand side of (32) equals 0. Note that since $A_e < A_{\max}$, the region in which $q^*_{EL} = 0$ (i.e., $A \in [A_e, \bar{A}]$) is larger than that in which $q^*_{IL} = 0$ (i.e., $A \in [A_{\max}, \bar{A}]$). We formally state these results in the following proposition.

Proposition 5 *Under the expected loss model, the bank makes the risk choice q^*_{EL} such that, if bank size $A \in [E, A_e)$, $q^*_{EL} \in (0, 1)$ and solves equation (32). Otherwise, for bank size $A \in [A_e, \bar{A}]$, $q^*_{EL} = 0$. The bank originates riskier loans ex ante under expected loss than incurred loss, i.e., $q^*_{EL} \leq q^*_{IL}$, where the inequality is strict if $A \in (A_{\min}, A_{\max})$. Ceteris paribus, the risk choice q^*_{EL} decreases in bank size A, increases in the face value of the low-risk loan α, and decreases in the face value of the high-risk loan β.*

Proposition 5 implies that when real effects are taken into account, there is a non trivial trade-off from adopting the expected loss model. While timely regulatory intervention provides discipline by curbing *ex post* inefficient asset substitution, banks respond to such intervention by originating riskier portfolios that, in turn, reduce surplus. We next investigate the conditions under which the *ex post* benefits of timely intervention exceed the *ex ante* costs triggered by such real effects and vice versa.

4.2.3 Surplus comparison with endogenous risk

To compare the results of endogenous risk choices with those of exogenous risk choices (Proposition 3), we first compare the surplus between the two provisioning models for a *given* bank size A.

Proposition 6 *With risk choices endogenized, there exists a set of thresholds $\left\{ \hat{A}, \hat{p}, \hat{s}_1 \right\}$, where $\hat{A} \in (A_{\min}, A_e)$ and $\hat{p} \in (0, 1)$, such that:*

1. *if $A \geq A_e$, the incurred loss model dominates the expected loss model;*

2. *if $A_{\min} < A < \hat{A}$ and $\Pr\left(s < \frac{1}{\alpha} | s_1 \leq \hat{s}_1\right) > \hat{p}$, the expected loss model dominates the incurred loss model.[13]*

Proposition 6 demonstrates that once we allow the bank's loan choices to respond to regulatory intervention, more timely information under the expected loss model is no longer always socially desirable. In fact, the expected loss model becomes inferior when banks are heavily leveraged. To see this, recall that when $A \geq A_e$, $q_{EL}^* = 0 < q_{IL}^*$. When the bank is highly leveraged, under the expected loss model, the bank always chooses the high-risk loan whereas under the incurred loss model, the bank chooses the low-risk loan with some probability, thus resulting in a higher surplus.

Proposition 6 also identifies sufficient conditions under which the benefits of *ex post* timely intervention under expected loss outweigh the costs triggered by its *ex ante* real effects. In particular, if the bank is not too highly leveraged *and* the low interim signal is sufficiently informative about the NPV of the low-risk project, then the expected loss

---

[13]Note that the bank never asset substitutes when $A \leq A_{\min}$ so the bank surplus is identical under both provisioning models.

model dominates the incurred loss model.[14] Intuitively, the bank is more likely to engage in asset substitution when the interim signal is low. Therefore, from a regulatory perspective, such early intervention entails the least amount of false alarms when the low interim signal is sufficiently informative. When this is the case, early intervention under expected loss generates a large benefit that dominates the cost stemming from the real effects. A key policy implication is that, once we endogenize the bank's risk choices, the efficiency of the expected loss model hinges on its accuracy in estimating early loan losses. Note that this implication stands in stark contrast to the case with exogenous risk choices in which, regardless of its precision, timely information under expected loss is always socially desirable. Stated differently, the accuracy–timeliness trade-off and the false alarm arguments proposed by opponents of expected loss models come to life *only* when one takes into account the real effects of accounting measurements.

Finally, the fact that the expected loss model improves the bank's surplus if leverage is sufficiently low but reduces the surplus otherwise suggests there exists a unique leverage threshold above which the expected loss model is dominated by the incurred loss model. Although the complexity of our model prevents us from showing the existence of such a unique threshold, numerical simulations suggest that such a conjecture is indeed true, as illustrated in Figure 4 which shows that $W_{EL} < W_{IL}$ if and only if the leverage A is large.

---

[14]More precisely, the condition $\Pr\left(s < \frac{1}{\alpha} | s_1 \leq \hat{s}_1\right) > \hat{p}$ requires that, conditional on some bad interim signals (i.e., $s_1 \leq \hat{s}_1$), the posterior belief that the low-risk project generates a negative NPV (i.e., $\alpha s < 1$) is sufficiently high. Furthermore, as shown in the proof of Proposition 6, under the linear-uniform structure of $s = ws_1 + (1 - w) s_2$, this condition requires that the precision w of s_1 is sufficiently large.
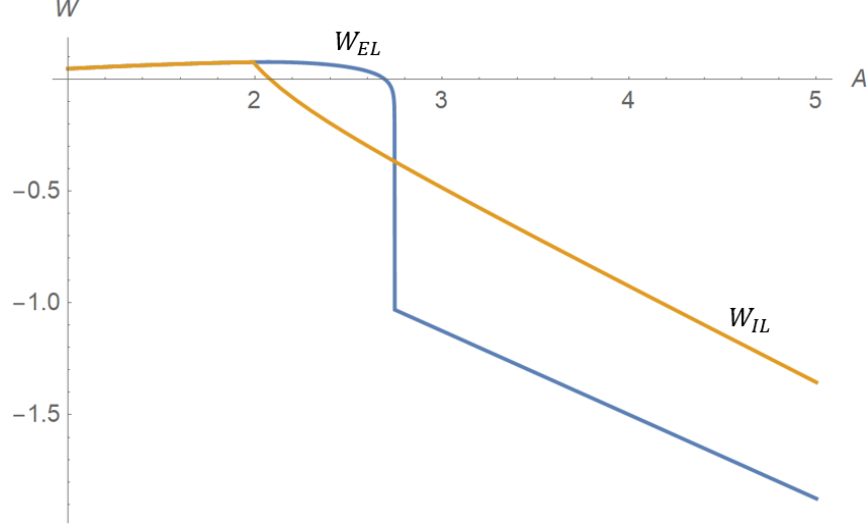
Figure 4: Surplus comparison with endogenous risk for $A \in [E, \bar{A}]$. The following parameter values are used in this plot: $\alpha = 1.5$, $\beta = 2.5$, $E = 1$, $\tau = 0.25$, $w = 0.45$, and $C(q) = \frac{q^{30}}{20}$.

4.2.4 Optimal design of capital requirements under CECL

An implication of Proposition 6 is that the regulator should take into account the real effects of loan loss provisioning models when designing the capital requirement policy. Put differently, changing the methodology for estimating loan losses requires the banking regulator to simultaneously adjust capital ratios. To better understand how the regulator should make such adjustments, we next solve for the optimal capital ratios under the two provisioning models.

We first reproduce the surplus under the incurred loss model, equation (25), here:

$$W_{IL}(A) = (1 - q_{IL}^*)A\tau\beta + q_{IL}^* \left(A \int_\tau^{\bar{s}_1(A)} \tau\beta g(s_1)ds_1 \right.$$
$$\left. + A \int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] \right) g(s_1)ds_1 \right) - A. \quad (36)$$

35

Taking the first order condition yields

$$
\begin{aligned}
\frac{\partial W_{IL}}{\partial A} \\
= \; & NPV_{IL}\left(\bar{s}_1(A)\right) - q_{IL}^* A \frac{\partial \bar{s}_1}{\partial A} \left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|\bar{s}_1(A)\right)ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1(A) \right] 1 - \tau\beta \right) g(\bar{s}_1(A)) \\
& + A \frac{\partial q_{IL}^*}{\partial A} \left(\int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] - \tau\beta \right) g(s_1)ds_1 \right). \quad (37)
\end{aligned}
$$

The first term in $\frac{\partial W_{IL}}{\partial A}$,

$$
\begin{aligned}
NPV_{IL}\left(\bar{s}_1(A)\right) \; = \; & \tau\beta - 1 + q_{IL}^* \left(\int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds \right] \alpha \right. \right. \\
& \left. \left. + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] - \tau\beta \right) g(s_1)ds_1 \right), \quad (38)
\end{aligned}
$$

measures the per-unit NPV from the bank's loan portfolio and represents the potential social benefit of increasing the bank's size. It is straightforward to verify the NPV is positive if and only if the asset size A is sufficiently small. The reason is that, if the bank is highly leveraged, it will convert its entire loan portfolio into high-risk in the interim, which, by Assumption 3, generates a negative NPV in expectation.

The other two terms in $\frac{\partial W_{IL}}{\partial A}$ are both negative and represent the social costs of increasing the bank's size. In particular, the second term captures the effect of increasing the asset size in motivating the bank to asset substitute *ex post* (i.e., $\frac{\partial \bar{s}_1}{\partial A} > 0$) whereas the third term captures the effect of increasing the asset size in discouraging the bank from exerting screening effort *ex ante* (i.e., $\frac{\partial q_{IL}^*}{\partial A} < 0$).[15]

---

[15]Mathematically, the second term in $\frac{\partial W_{IL}}{\partial A}$ is negative because 1) from Proposition 1, $\frac{\partial \bar{s}_1}{\partial A} > 0$ and 2) $\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|\bar{s}_1(A)\right)ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1(A) \right] > 1 > \tau\beta$ (Assumption 3). The third term is negative because, 1)

The regulator sets the optimal capital requirement ratio by trading off the above benefit against the costs. We denote the optimal bank size under the incurred loss model by A_{IL}^* which solves the first order condition (37). The optimal bank size under the expected loss model can be similarly derived. We denote it by A_{EL}^*.[16] Because A_{IL}^* and A_{EL}^* are defined by implicit solutions to differential equations, in general, we are unable to compare them analytically. To facilitate this comparison, we hereafter impose two additional assumptions for the rest of the analysis.

Assumption 5: The surplus functions W_{IL} and W_{EL} are both single-peaked in $A \in \left[E, \bar{A}\right]$, i.e., there exists a unique A_i^*, where $i \in \{IL, EL\}$, such that $\frac{\partial W_i}{\partial A} > 0$ if and only if $A < A_i^*$.

The assumption guarantees the uniqueness of the optimal capital requirements and rules out scenarios in which there are multiple spikes in the surplus functions. Numerical examples suggest that the surplus functions are indeed single-peaked, as illustrated in Figure 4.

Assumption 6: Absent asset substitution, the bank's loan portfolio, *ex ante*, generates a positive NPV, i.e., for $A \le A_{\min}$, $W_i \ge 0$, where $i \in \{IL, EL\}$.

The assumption rules out uninteresting scenarios in which the cost of screening $C(q)$ is so steep that the bank's incentives to screen borrowers and originate low-risk loans are low even when the bank's leverage is relatively low.

Given these two assumptions, we next derive a sufficient condition under which $A_{EL}^* > A_{IL}^*$, i.e., the regulator should lower the capital ratio in response to the adoption of the expected loss model.

---

from Proposition 4, $\frac{\partial q_{IL}^*}{\partial A} < 0$ and 2) $\left[\int_{\frac{1}{\alpha}}^{1} s f\left(s|s_1\right) ds\right] \alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right] > 1 > \tau\beta$.

[16] The first order condition that defines A_{EL}^* is complicated and we derive it in the proof of Proposition 7.

Proposition 7 *There exist some thresholds $\left\{\underline{\beta}, \underline{s}_1, \underline{p}, \underline{g}\right\}$, such that, if $\Pr\left(s < \frac{1}{\alpha} | s_1 \leq \underline{s}_1\right) > \underline{p}$, $\inf_{s_1 \leq \underline{s}_1} g(s_1) > \underline{g}$, and $\beta < \underline{\beta}$, the regulator sets tighter capital requirements under the incurred loss model than under the expected loss model, i.e., $A_{IL}^* < A_{EL}^*$, and the expected loss model dominates the incurred loss model, i.e., $W_{EL}(A_{EL}^*) > W_{IL}(A_{IL}^*)$.*

Proposition 7 identifies a region of parameters in which the regulator may be able to relax the capital requirements when banks use the expected loss model for loan loss provisioning. In that region, both the precision and the likelihood of the bad interim information are relatively high, and/or when the incentives to engage in inefficient asset substitution are not too severe. Intuitively, the likelihood of asset substitution is greater if the interim signal is more likely to be bad. In this case, when the expected loss model provides more precise interim information, it reduces the amount of false alarm cost from early intervention, thus making the intervention more effective in curbing *ex post* asset-substitution. Since the regulation has become more effective under expected loss, the regulator can actually relax the capital requirement.

Note that the conditions in Proposition 7 are *sufficient* conditions for relaxing capital requirements under expected loss. When the conditions are not met, it is difficult to compare the optimal capital requirements across the loan loss models analytically, and we thus turn to numerical analyses. Using the linear-uniform structure, Figure 5 provides a numerical example that illustrates the comparison between the optimal capital requirements. The figure shows that the regulator should loosen the capital requirement under the expected loss model in the lower-right region when the expected loss model reports more precise interim information (high w) and/or the high-risk loan is less attractive to the bank (low
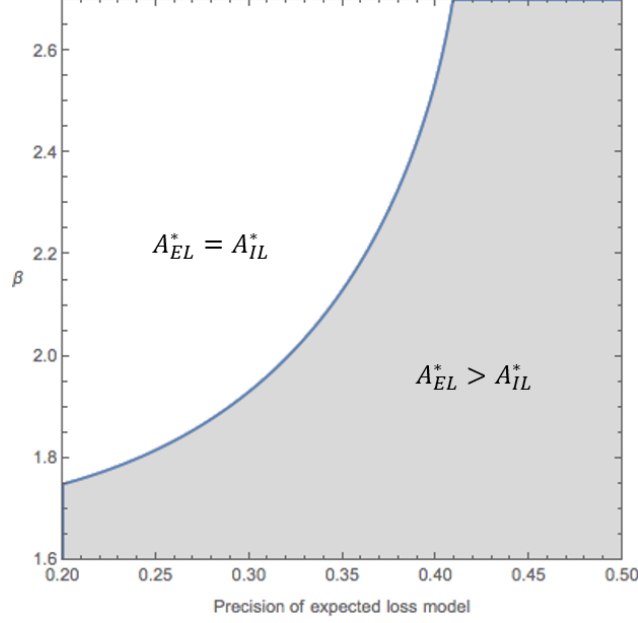
Figure 5: Comparison between optimal capital requirements under the expected loss model and the incurred loss model. The following parameter values are used in this plot: $\alpha = 1.5$, $E = 1$, $\tau = 0.25$ and $C(q) = \frac{q^{30}}{20}$.

β). On the other hand, when the bank finds the high-risk loan more appealing to invest in (high β), the excessive risk-taking problem becomes sufficiently severe that the regulator sets an extremely stringent capital requirement under any loan loss provisioning models, i.e., both A_{EL}^* and A_{IL}^* fall below $A_{\min}$, such that the surplus under the two loan loss models overlaps. In this case, the optimal capital requirements are the same across the two loan loss provisioning models.

Another key message of Proposition 7 is that, provided that the capital requirements are optimally set, the expected loss model generates a higher surplus than the incurred loss model. Stated differently, while Proposition 6 suggests that, fixing the capital requirement policy, the timely information released under expected loss, if not sufficiently precise, may impair surplus, Proposition 7 shows that such adverse effect can be overturned if the regulator can appropriately tailor the capital ratio in response to the provisioning model. These results

39

echo the recent call for better coordination between accounting and bank regulation.[17]

5   Conclusion

The loan loss provision is arguably the largest accrual item on banks' financial statements and thus "plays a prominent role in much of the bank accounting literature" (Beatty and Liao, 2014). Yet, somewhat surprisingly, relatively little is known about the exact mechanism through which loan loss provisioning interacts with prudential regulation to affect banks' behavior. We believe that, to better inform policy debates, it is crucial to have a solid conceptual understanding of the role of loan loss provisioning and how it interacts with bank capital. An important insight of our study is that relying on expected-loss models such as CECL would curb lending only if banks' capital requirements are set independently of the accounting standards used to provision for loan losses. But if bank regulators optimally use balance-sheet information to tailor banks' capital requirements to the riskiness of banks' loan portfolios, they could potentially relax capital requirements to spur lending.

Our model may be used as a springboard to study other important aspects of loan loss provisioning that we did not capture in our environment. Banks are often criticized for exercising a large amount of discretion in estimating and disclosing their loan losses, even under the incurred loss model. Arguably, such discretion will only increase when it comes to estimating expected loan losses. To the extent that banks' discretionary reporting of loan loss provisions hides loan loss information from regulators, it may result in an additional

---

[17]The U.S. Congress recently recognized the importance of adjusting capital requirements in light of banks implementing CECL. It has directed the U.S. Treasury Department, in consultation with bank regulators, to study the impact of the CECL and to determine whether any changes to regulatory capital requirements are necessary. (Maurer, 2020).

cost of the expected loss model that we have ignored in our analysis. How such a cost of reporting discretion may be traded against the cost and the benefit identified in our model is an interesting avenue we leave to future research.

We have also not considered how loan loss provisioning may lead to spillovers and affect the systemic risk of the banking industry (e.g., procyclicality of bank lending). In the current debate, there are concerns that the incurred loss model may have contributed to procyclicality and adopting the new expected loss model may help to mitigate procyclicality.[18] Extending our single-bank model to include multiple banks may thus shed light on the systemic impact of loan loss provisioning, which is another interesting avenue left for future research.

---

[18]For theoretical and empirical work on provisioning and procyclicality, see Abad and Suarez (2018), Agenor and Zilberman (2015), Bouvatier and Lepetit (2012), Dewatripont and Tirole (2012) and Goncharenko and Rauf (2019).

References

[1] Abad, Jorge, and Javier Suarez. "The procyclicality of expected credit loss provisions." (2018).

[2] Acharya, Viral V., and Stephen G. Ryan. "Banks' financial reporting and financial system stability." Journal of Accounting Research 54, no. 2 (2016): 277-340.

[3] Acharya, Viral V., and Tanju Yorulmazer. "Too many to fail—An analysis of time-inconsistency in bank closure policies." Journal of Financial Intermediation 16, no. 1 (2007): 1-31.

[4] Agenor, Pierre-Richard, and Roy Zilberman. "Loan loss provisioning rules, procyclicality, and financial volatility." Journal of Banking and Finance 61 (2015): 301-315.

[5] Allen, Franklin, and Elena Carletti. "Mark-to-market accounting and liquidity pricing." Journal of Accounting and Economics 45, no. 2-3 (2008): 358-378.

[6] Akins, Brian, Yiwei Dou, and Jeffrey Ng. "Corruption in bank lending: The role of timely loan loss recognition." Journal of Accounting and Economics 63, no. 2-3 (2017): 454-478.

[7] Bagehot, Walter. Lombard Street: A description of the money market. London: HS King, 1873.

[8] Basel Committee. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." Basel Committee on Banking Supervision, Basel (2010).

[9] Beatty, Anne, and Scott Liao. "Do delays in expected loss recognition affect banks' willingness to lend?" Journal of Accounting and Economics 52, no. 1 (2011): 1-20.

[10] Beatty, Anne, and Scott Liao. "Financial accounting in the banking industry: A review of the empirical literature." Journal of Accounting and Economics 58, no. 2-3 (2014): 339-383.

[11] Bertomeu, Jeremy, Lucas Mahieux, and Haresh Sapra. "Interplay between accounting and prudential regulation." Available at SSRN 3266348 (2020).

[12] Bouvatier, Vincent, and Laetitia Lepetit. "Provisioning rules and bank lending: A theoretical model." Journal of Financial Stability 8, no. 1 (2012): 25-31.

[13] Burkhardt, Katrin, and Roland Strausz. "Accounting transparency and the asset substitution problem." The Accounting Review 84, no. 3 (2009): 689-712.

[14] Bushman, Robert M., and Christopher D. Williams. "Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking." Journal of Accounting and Economics 54, no. 1 (2012): 1-18.

[15] Bushman, Robert M., and Christopher D. Williams. "Delayed expected loss recognition and the risk profile of banks." Journal of Accounting Research 53, no. 3 (2015): 511-553.

[16] Corona, Carlos, Lin Nan, and Gaoqing Zhang. "Accounting Information Quality, Interbank Competition, and Bank Risk Taking." The Accounting Review 90 (2015): 967-985.

[17] Corona, Carlos, Lin Nan, and Gaoqing Zhang. "Banks' asset reporting frequency and capital regulation: An analysis of discretionary use of fair-value accounting." The Accounting Review 94, no. 2 (2019): 157-178.

[18] Corona, Carlos, Lin Nan, and Gaoqing Zhang. "The Coordination Role of Stress Tests in Bank Risk Taking." Journal of Accounting Research 57, no. 5 (2019): 1161-1200.

[19] Dewatripont, Mathias, and Jean Tirole. The prudential regulation of banks. Vol. 1. MIT press, 1994.

[20] Dewatripont, Mathias, and Jean Tirole. "Macroeconomic shocks and bank regulation." Journal

of Money, Credit and Banking 44 (2012): 237-254.

[21] Financial Accounting Standards Board. Accounting Standards Update No. 2016-13 Financial Instruments-Credit Losses (Topic 326), June, 2016.

[22] Financial Accounting Standards Board. Public round table meeting on credit losses, January 28, 2019, available at https://www.youtube.com/watch?v=FH7xWZsCX9s&t=2866s.

[23] Freixas, Xavier. "Optimal bail out policy, conditionality and constructive ambiguity." (1999).

[24] Gao, Pingyang, and Xu Jiang. "Reporting choices in the shadow of bank runs." Journal of Accounting and Economics 65.1 (2018): 85-108.

[25] Goldstein, Itay, and Haresh Sapra. "Should banks' stress test results be disclosed? An analysis of the costs and benefits." Foundations and Trends in Finance 8, no. 1 (2014): 1-54.

[26] Goncharenko, Roman, and Asad Rauf. "Loan Loss Provisioning Requirements in a Dynamic Model of Banking." Available at SSRN 3247631 (2019).

[27] Heaton, John C., Deborah Lucas, and Robert L. McDonald. "Is mark-to-market accounting destabilizing? Analysis and implications for policy." Journal of Monetary Economics 57, no. 1 (2010): 64-75.

[28] Laux, Christian, and Thomas Rauter. "Procyclicality of US bank leverage." Journal of Accounting Research 55, no. 2 (2017): 237-273.

[29] Liang, Pierre Jinghong and Gaoqing Zhang. "On the Social Value of Accounting Objectivity in Financial Stability." The Accounting Review 94, no 1 (2019): 229-248.

[30] Lu, Tong, Haresh Sapra, and Ajay Subramanian. "Agency Conflicts, Bank Capital Regulation, and Marking-to-Market." The Accounting Review 94, no 6 (2019): 365-384.

[31] Mahieux, Lucas. "Fair Value Accounting, Transparency, and Financial Stability." Working paper, September, 2019.

[32] Mailath, George J., and Loretta J. Mester. "A positive analysis of bank closure." Journal of Financial Intermediation 3, no. 3 (1994): 272-299.

[33] Maurer, Mark. "Lawmakers Press Rule Maker on Economic Impact of Credit Loss Standard." Wall Street Journal, 2020, available at https://www.wsj.com/articles/lawmakers-press-rule-maker-on-economic-impact-of-credit-loss-standard-11579139500.

[34] Plantin, Guillaume, Haresh Sapra, and Hyun Song Shin. "Marking-to-market: panacea or Pandora's box?." Journal of Accounting Research 46, no. 2 (2008): 435-460.

[35] Quaadman, Tom. U.S. Chamber of Commerce's comment letter to the SEC and FASB, January, 2019, available at https://www.centerforcapitalmarkets.com/wp-content/uploads/2019/01/Chamber-CECL-Letter.pdf?#.

[36] Zhang, Gaoqing. "Competition and Opacity in the Financial System." Forthcoming, Management Science (2019).

Appendix: proofs

Proof. of Proposition 1: We first derive the threshold on asset substitution $\bar{s}_1$. At one extreme, $A = E$ (the lower bound for A),

$$\tau\beta < 1 \leq \left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1 = \tau) \, ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right]. \tag{39}$$

That is, at $A = E$, the bank never engages in asset substitution, i.e., $\bar{s}_1 = \tau$. At the other extreme of $A = \bar{A}$, by Assumption 4, the bank always asset-substitutes, i.e., $\bar{s}_1 = 1$. In addition, note that the left hand side of (19) is increasing at a faster speed in A than its right hand side, i.e.,

$$\tau(\beta - 1) > \left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1) \, ds\right](\alpha - 1). \tag{40}$$

The inequality is by Assumption 4. Therefore, there exist some $A_{\min} > E$ and $A_{\max} \in (A_{\min}, \bar{A})$, such that: for $A \in (A_{\min}, A_{\max})$, $\bar{s}_1(A) \in (\tau, 1)$; for $A \leq A_{\min}$, $\bar{s}_1(A) = \tau$; for $A \geq A_{\max}$, $\bar{s}_1(A) = 1$. The cutoff $A_{\max}$ is defined such that

$$\tau = \left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1 = 1) \, ds\right]\frac{A_{\max}\alpha - (A_{\max} - E)}{A_{\max}\beta - (A_{\max} - E)} + \Pr\left[s < \frac{1}{\alpha}|s_1 = 1\right]\frac{E}{A_{\max}\beta - (A_{\max} - E)}, \tag{41}$$

The cutoff $A_{\min}$ is defined such that

$$\tau = \left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1 = \tau) \, ds\right]\frac{A_{\min}\alpha - (A_{\min} - E)}{A_{\min}\beta - (A_{\min} - E)} + \Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right]\frac{E}{A_{\min}\beta - (A_{\min} - E)}. \tag{42}$$

Next, we derive some comparative statics on $\bar{s}_1 \in (\tau, 1)$, which is defined such that

$$\tau = \left[\int_{\frac{1}{\alpha}}^{1} sf(s|\bar{s}_1) \, ds \right] \frac{A\alpha - (A - E)}{A\beta - (A - E)} + \Pr \left[s < \frac{1}{\alpha} \Big| \bar{s}_1 \right] \frac{E}{A\beta - (A - E)}, \tag{43}$$

and is equivalent to

$$\tau \left(A\beta - (A - E) \right) = \left[\int_{\frac{1}{\alpha}}^{1} sf(s|\bar{s}_1) \, ds \right] \left(A\alpha - (A - E) \right) + \Pr \left[s < \frac{1}{\alpha} \Big| \bar{s}_1 \right] E. \tag{44}$$

Taking the derivative of (44) with respect to A, we get

$$\tau(\beta - 1) = \left[\frac{\partial \bar{s}_1}{\partial A} \int_{\frac{1}{\alpha}}^{1} s \frac{\partial f(s|\bar{s}_1)}{\partial s_1} \, ds \right] \left(A\alpha - (A - E) \right) \tag{45}$$

$$+ \left[\frac{\partial \bar{s}_1}{\partial A} \int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f(s|\bar{s}_1)}{\partial s_1} \, ds \right] E + \left[\int_{\frac{1}{\alpha}}^{1} sf(s|\bar{s}_1) \, ds \right] (\alpha - 1), \tag{46}$$

which is equivalent to

$$\frac{\partial \bar{s}_1}{\partial A} = \frac{\tau(\beta - 1) - \left[\int_{\frac{1}{\alpha}}^{1} sf(s|\bar{s}_1) \, ds \right] (\alpha - 1)}{\left[\int_{\frac{1}{\alpha}}^{1} s \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds \right] \left(A\alpha - (A - E) \right) + \left[\int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds \right] E} \geq 0. \tag{47}$$

The numerator of $\frac{\partial \bar{s}_1}{\partial A}$ is positive by Assumption 4. The denominator of $\frac{\partial \bar{s}_1}{\partial A}$ is positive because

the right hand side of (44) increases in s_1, i.e., a higher s_1 improves the bank's expected

payoff from the low-risk loan. To see this, rewriting the right hand side of (44) gives that:

$$\left[\int_{\frac{1}{\alpha}}^{1} sf\left(s|\bar{s}_1\right) ds\right] \left(A\alpha - (A - E)\right) + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1\right] E \tag{48}$$

$$= A\alpha - (A - E) - \left(A - \frac{A - E}{\alpha}\right) F\left(\frac{1}{\alpha}|\bar{s}_1\right) - \left[\int_{\frac{1}{\alpha}}^{1} F\left(s|\bar{s}_1\right) ds\right] \left(A\alpha - (A - E)\right)$$

$$+ F\left(\frac{1}{\alpha}|\bar{s}_1\right) E$$

$$= \left[A\alpha - (A - E)\right] \left[1 - \int_{\frac{1}{\alpha}}^{1} F\left(s|\bar{s}_1\right) ds\right] - \left(1 + \frac{1}{\alpha}\right) F\left(\frac{1}{\alpha}|\bar{s}_1\right).$$

The first step uses integration by parts. Taking the derivative of (44) with respect to $\bar{s}_1$ gives that

$$-\left[A\alpha - (A - E)\right] \left[\int_{\frac{1}{\alpha}}^{1} \frac{\partial F\left(s|\bar{s}_1\right)}{\partial s_1} ds\right] - \left(1 + \frac{1}{\alpha}\right) \frac{\partial F\left(\frac{1}{\alpha}|\bar{s}_1\right)}{\partial s_1} > 0. \tag{49}$$

The inequality uses the property that s_1 improves the posterior distribution of s in the sense of first-order stochastic dominance, i.e., $\frac{\partial F(s|\bar{s}_1)}{\partial s_1} < 0$. Therefore, the right hand side of (44) increases in s_1, which implies that:

$$\left[\int_{\frac{1}{\alpha}}^{1} s\frac{\partial f\left(s|\bar{s}_1\right)}{\partial s_1} ds\right] \left(A\alpha - (A - E)\right) + \left[\int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f\left(s|\bar{s}_1\right)}{\partial s_1} ds\right] E > 0. \tag{50}$$

Taking the derivative of (44) with respect to β, we get

$$\tau A = \left[\frac{\partial \bar{s}_1}{\partial \beta} \int_{\frac{1}{\alpha}}^{1} s\frac{\partial f\left(s|\bar{s}_1\right)}{\partial s_1} ds\right] \left(A\alpha - (A - E)\right) + \left[\frac{\partial \bar{s}_1}{\partial \beta} \int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f\left(s|\bar{s}_1\right)}{\partial s_1} ds\right] E, \tag{51}$$

which is equivalent to

$$\frac{\partial \bar{s}_1}{\partial \beta} = \frac{\tau A}{\left[\int_{\frac{1}{\alpha}}^{1} s \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds\right](A\alpha - (A - E)) + \left[\int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds\right] E} \geq 0. \tag{52}$$

Finally, taking the derivative of (44) with respect to α, we get

$$0 = \left[\frac{\partial \bar{s}_1}{\partial \alpha} \int_{\frac{1}{\alpha}}^{1} s \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds + \frac{1}{\alpha^3} f\left(\frac{1}{\alpha}|\bar{s}_1\right)\right](A\alpha - (A - E)) + \left[\int_{\frac{1}{\alpha}}^{1} sf(s|\bar{s}_1) ds\right] A$$
$$+ \left[\frac{\partial \bar{s}_1}{\partial \alpha} \int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds - \frac{1}{\alpha^2} f\left(\frac{1}{\alpha}|\bar{s}_1\right)\right] E, \tag{53}$$

which is equivalent to

$$\frac{\partial \bar{s}_1}{\partial \alpha} = \frac{-\frac{1}{\alpha^3} f\left(\frac{1}{\alpha}|\bar{s}_1\right)(\alpha - 1)(A - E) - \left[\int_{\frac{1}{\alpha}}^{1} sf(s|\bar{s}_1) ds\right] A}{\left[\int_{\frac{1}{\alpha}}^{1} s \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds\right](A\alpha - (A - E)) + \left[\int_{\tau}^{\frac{1}{\alpha}} \frac{\partial f(s|\bar{s}_1)}{\partial s_1} ds\right] E} \leq 0. \tag{54}$$

∎

Proof. of Proposition 2: See the main text. ∎

Proof. of Proposition 3: See the main text. ∎

Proof. of Proposition 4: From the main text, for $A \in [E, A_{\max})$, q_{IL}^* solves

$$\int_{\bar{s}_1(A)}^{1} \left(\left[\int_{\frac{1}{\alpha}}^{1} sf(s|s_1) ds\right](A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|s_1\right] E\right) g(s_1) ds_1 \tag{55}$$
$$- \int_{\bar{s}_1(A)}^{1} \tau(A\beta - (A - E)) g(s_1) ds_1$$
$$= C'(q_{IL}^*).$$

Using the implicit function theorem,

$$
\begin{aligned}
C''(q_{IL}^*)\frac{\partial q_{IL}^*}{\partial A} &= \int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds \right] (\alpha - 1) - \tau(\beta - 1) \right) g(s_1)ds_1 \\
&\quad -\frac{\partial \bar{s}_1}{\partial A}\left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|\bar{s}_1(A))\,ds \right] (A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1(A) \right] E \right. \\
&\quad \left. -\tau(A\beta - (A - E)) \right) g(\bar{s}_1(A)) \\
&= \int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds \right] (\alpha - 1) - \tau(\beta - 1) \right) g(s_1)ds_1 \qquad (56) \\
&\leq 0.
\end{aligned}
$$

The second step uses the definition of $\bar{s}_1(A)$ in (44):

$$
\left[\int_{\frac{1}{\alpha}}^1 sf(s|\bar{s}_1(A))\,ds \right] (A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1(A) \right] E = \tau(A\beta - (A - E)). \qquad (57)
$$

This last step uses Assumption 4:

$$
\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds \right] (\alpha - 1) < \tau(\beta - 1). \qquad (58)
$$

The inequality is strict if and only if $\bar{s}_1(A) < 1$ (i.e., $A < A_{\max}$).

Taking the derivative of (55) with respect to α,

$$
C''(q_{IL}^*)\frac{\partial q_{IL}^*}{\partial \alpha} \tag{59}
$$

$$
= \int_{\bar{s}_1(A)}^1 \frac{\partial}{\partial \alpha}\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right](A\alpha - (A - E)) + \Pr\left[s < \frac{1}{\alpha}|s_1\right]E\right)g(s_1)ds_1
$$

$$
+ \frac{\partial \bar{s}_1}{\partial \alpha}\left(-\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|\bar{s}_1\right)ds\right](A\alpha - (A - E)) - \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1\right]E + \tau\left(A\beta - (A - E)\right)\right)g(\bar{s}_1)
$$

$$
= \int_{\bar{s}_1(A)}^1\left[\frac{1}{\alpha^3}f\left(\frac{1}{\alpha}|s_1\right)(\alpha - 1)(A - E) + \left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right]A\right]g(s_1)ds_1
$$

$$
\geq 0.
$$

The second step uses (57). The last inequality is strict if and only if $\bar{s}_1(A) < 1$.

Lastly, taking the derivative of (55) with respect to β,

$$
C''(q_{IL}^*)\frac{\partial q_{IL}^*}{\partial \beta} \tag{60}
$$

$$
= -\int_{\bar{s}_1(A)}^1 \tau A g(s_1)ds_1
$$

$$
+ \frac{\partial \bar{s}_1}{\partial \beta}\left(-\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right](A\alpha - (A - E)) - \Pr\left[s < \frac{1}{\alpha}|s_1\right]E + \tau\left(A\beta - (A - E)\right)\right)g(\bar{s}_1)
$$

$$
= -\int_{\bar{s}_1(A)}^1 \tau A g(s_1)ds_1
$$

$$
\leq 0.
$$

The second step uses (57). The last inequality is strict if and only if $\bar{s}_1(A) < 1$. ∎

Proof. of Proposition 5: We supplement the proofs in the main text by computing the comparative statics regarding q_{EL}^*. For our convenience, we first reproduce the first order

condition (32) on q^*_{EL} below:

$$C'(q^*_{IL}) + \int_\tau^{\bar{s}_1(A)} [E - \tau (A\beta - (A - E))] \, g(s_1) \, ds_1 = C'(q^*_{EL}). \tag{61}$$

We first derive the comparative statics regarding A. Taking the derivative of (61) with respect to A,

$$
\begin{aligned}
& C''(q^*_{EL}) \frac{\partial q^*_{EL}}{\partial A} \\
= \ & C''(q^*_{IL}) \frac{\partial q^*_{IL}}{\partial A} - G(\bar{s}_1(A)) \tau(\beta - 1) + \frac{\partial \bar{s}_1}{\partial A}[E - \tau(A\beta - (A - E))]g(\bar{s}_1(A)) \leq 0. \tag{62}
\end{aligned}
$$

Note that the first term of (62) is non-positive because from Proposition 4, $\frac{\partial q^*_{IL}}{\partial A} \leq 0$. The second term of (62) is non-positive because $-\tau(\beta - 1) < 0$. Finally, the third term of (62) is also non-positive. This is because if $A \leq A_{\min}$, from Proposition 1, $\bar{s}_1(A) = \tau$ for all $A \leq A_{\min}$, which makes the third term of (62) equal zero. If $A > A_{\min}$, the third term of (62) is also non-positive because $\frac{\partial \bar{s}_1}{\partial A} \geq 0$ and $E - \tau(A\beta - (A - E)) < 0$ given (33) in the main text. Since all the terms of (62) are non-positive, $\frac{\partial q^*_{EL}}{\partial A} \leq 0$.

Next, we derive the comparative statics regarding α. Taking the derivative of (61) with respect to α,

$$C''(q^*_{EL}) \frac{\partial q^*_{EL}}{\partial \alpha} = C''(q^*_{IL}) \frac{\partial q^*_{IL}}{\partial \alpha} + \frac{\partial \bar{s}_1}{\partial \alpha}[E - \tau(A\beta - (A - E))]g(\bar{s}_1(A)) \geq 0. \tag{63}$$

The last inequality uses 1) from Proposition 4, $\frac{\partial q^*_{IL}}{\partial \alpha} \geq 0$ and 2) $\frac{\partial \bar{s}_1}{\partial \alpha} \leq 0$ such that $\frac{\partial \bar{s}_1}{\partial \alpha}[E - \tau(A\beta - (A - E))] \geq 0$.

Finally, we derive the comparative statics regarding β. Taking the derivative of (61) with respect to β,

$$C''(q^*_{EL})\frac{\partial q^*_{EL}}{\partial \beta} = C''(q^*_{IL})\frac{\partial q^*_{IL}}{\partial \beta} + \frac{\partial \bar{s}_1}{\partial \beta}[E - \tau(A\beta - (A - E))]g(\bar{s}_1(A)) - G(\bar{s}_1(A))\tau A \leq 0. \quad (64)$$

The last inequality uses 1) from Proposition 4, $\frac{\partial q^*_{IL}}{\partial \beta} \leq 0$ and 2) $\frac{\partial \bar{s}_1}{\partial \beta} \geq 0$ such that $\frac{\partial \bar{s}_1}{\partial \beta}[E - \tau(A\beta - (A - E))] \leq 0$. ∎

Proof. of Proposition 6: From (25) and (26) of the main text, the surplus is given by

$$W_{IL}(A) = (1 - q^*_{IL})A\tau\beta + q^*_{IL}\left(A\int_\tau^{\bar{s}_1(A)} \tau\beta g(s_1)ds_1 \right.$$

$$\left. + A\int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1\right) - A,$$

and

$$W_{EL}(A) = (1 - q^*_{EL})A\tau\beta + q^*_{EL}\left(A\int_\tau^{\bar{s}_1(A)} g(s_1)ds_1 \right.$$

$$\left. + A\int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1\right) - A.$$

Taking the difference gives that:

$$W_{EL}(A) - W_{IL}(A) \tag{65}$$

$$= A\int_\tau^{\bar{s}_1(A)} (q^*_{EL} - q^*_{IL}\tau\beta)\,g(s_1)ds_1$$

$$+ (q^*_{EL} - q^*_{IL})A\left(\int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right).$$

First, consider the case that $A \geq A_e$. From Proposition 4 and Proposition 5, $q_{IL}^* \geq 0$ and $q_{EL}^* = 0$, where the inequality is strict if and only if $A < A_{\max}$. Plugging $q_{EL}^* = 0$ into the expression of $W_{EL}(A) - W_{IL}(A)$ gives that:

$$W_{EL}(A) - W_{IL}(A) \tag{66}$$
$$= -q_{IL}^* \left(A \int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] - \tau\beta \right) g(s_1)ds_1 \right)$$
$$\leq 0.$$

The last step uses that $\tau\beta < 1$. The inequality is strict if and only if $q_{IL}^* > 0$, which holds if and only if $A < A_{\max}$.

Second, consider the case that $A \leq A_{\min}$:

$$W_{EL}(A) - W_{IL}(A) \tag{67}$$
$$= A \int_{\tau}^{\bar{s}_1(A)} (q_{EL}^* - q_{IL}^*\tau\beta)\, g(s_1)ds_1$$
$$+ (q_{EL}^* - q_{IL}^*)\, A\left(\int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha}|s_1 \right] \right) g(s_1)ds_1 - \tau\beta \right)$$
$$= 0. \tag{68}$$

The last equality uses that, from Proposition 5, for $A \leq A_{\min}$, $\bar{s}_1 = \tau$ and $q_{EL}^* = q_{IL}^* > 0$.

Finally, consider a small neighborhood of $A \in (A_{\min}, A_{\min} + \varepsilon)$, where $\varepsilon > 0$ is arbitrarily

small. If $\Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right] = 1$, we have

$$W_{EL}(A) - W_{IL}(A)$$

$$= W_{EL}(A_{\min}) - W_{IL}(A_{\min}) + \varepsilon \lim_{A \to A_{\min}^+} \left(\frac{\partial W_{EL}}{\partial A} - \frac{\partial W_{IL}}{\partial A}\right)$$

$$\propto \lim_{A \to A_{\min}^+} \left(\frac{\partial W_{EL}}{\partial A} - \frac{\partial W_{IL}}{\partial A}\right)$$

$$= \frac{\partial q_{EL}^*}{\partial A}\left(A_{\min}\int_\tau^{\bar{s}_1} g(s_1)ds_1 + A_{\min}\int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - A_{\min}\tau\beta\right)$$

$$- \frac{\partial q_{IL}^*}{\partial A}\left(A_{\min}\int_\tau^{\bar{s}_1} \tau\beta g(s_1)ds_1 + A_{\min}\int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1\right.$$

$$\left. - A_{\min}\tau\beta\right) + q_{EL}^*\left(\int_\tau^{\bar{s}_1} g(s_1)ds_1 + \int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right)$$

$$- q_{IL}^*\left(\int_\tau^{\bar{s}_1} \tau\beta g(s_1)ds_1 + \int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right)$$

$$+ q_{EL}^* \frac{\partial \bar{s}_1}{\partial A} A_{\min} g(\bar{s}_1)\left(1 - \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|\bar{s}_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1\right]\right)\right)$$

$$- q_{IL}^* \frac{\partial \bar{s}_1}{\partial A} A_{\min} g(\bar{s}_1)\left(\tau\beta - \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|\bar{s}_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|\bar{s}_1\right]\right)\right)$$

$$= A_{\min} \lim_{A \to A_{\min}^+} \left(\frac{\partial q_{EL}^*}{\partial A} - \frac{\partial q_{IL}^*}{\partial A}\right)\left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right)$$

$$+ q_{EL}^* \frac{\partial \bar{s}_1}{\partial A} A_{\min} g(\bar{s}_1)(1 - \tau\beta)$$

$$= A_{\min} g(\tau)\frac{\partial \bar{s}_1}{\partial A}\left(q_{EL}^*(1 - \tau\beta)\right.$$

$$+ \frac{[E - \tau(A_{\min}\beta - (A_{\min} - E))]}{C''(q_{EL}^*)}\left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right)\right)$$

$$= A_{\min} g(\tau)\frac{\partial \bar{s}_1}{\partial A}q_{EL}^*(1 - \tau\beta) > 0.$$

"$\propto$" stands that "having the same sign as." The second step uses $W_{EL}(A_{\min}) = W_{IL}(A_{\min})$.

The fourth step uses $\bar{s}_1(A_{\min}) = \tau$ and $q_{EL}^*(A_{\min}) = q_{IL}^*(A_{\min})$. The fifth step uses that,

from (62),

$$C''(q^*_{EL})\frac{\partial q^*_{EL}}{\partial A} - C''(q^*_{IL})\frac{\partial q^*_{IL}}{\partial A} = \int_\tau^{\bar{s}_1(A)} [-\tau(\beta-1)]g(s_1)ds_1 + \frac{\partial \bar{s}_1}{\partial A}[E - \tau(A\beta - (A-E))]g(\bar{s}_1(A)).$$

(69)

At $A = A_{\min}$, this equation can be reduced into:

$$C''(q^*_{EL}) \lim_{A \to A^+_{\min}} \left(\frac{\partial q^*_{EL}}{\partial A} - \frac{\partial q^*_{IL}}{\partial A} \right) = \frac{\partial \bar{s}_1}{\partial A}[E - \tau(A\beta - (A-E))]g(\bar{s}_1(A)). \qquad (70)$$

The last step uses (42):

$$\tau = \left[\int_{\frac{1}{\alpha}}^1 s f(s|s_1 = \tau) ds \right] \frac{A_{\min}\alpha - (A_{\min} - E)}{A_{\min}\beta - (A_{\min} - E)} + \Pr\left[s < \frac{1}{\alpha} | s_1 = \tau \right] \frac{E}{A_{\min}\beta - (A_{\min} - E)}.$$

(71)

If $\Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right] = 1$, $\tau[A_{\min}\beta - (A_{\min} - E)] = E$. Under the linear-uniform structure of $s = ws_1 + (1-w)s_2$, $\Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right] = 1$ requires that at $s_1 = \tau$ and $s_2 = 1$,

$$s = w\tau + (1-w) < \frac{1}{\alpha}, \qquad (72)$$

which reduces into $w > \frac{\alpha-1}{\alpha-\alpha\tau}$. $\frac{\alpha-1}{\alpha-\alpha\tau} < 1$ because $\alpha\tau < \beta\tau < 1$.

Since in a small neighborhood of $A \in (A_{\min}, A_{\min} + \varepsilon)$, $W_{EL} > W_{IL}$ if $\Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right] = 1$. By continuity, there exists some thresholds $\left\{\hat{p}, \hat{A}, \hat{s}_1\right\}$ that satisfy $A_{\min} < \hat{A} < A_e$, such that for $\Pr\left[s < \frac{1}{\alpha}|s_1 \le \hat{s}_1\right] > \hat{p}$ and $A < \hat{A}$, $W_{EL} \ge W_{IL}$. $\hat{A} < A_e$ because at $A = A_e$, $W_{EL} < W_{IL}$. ∎

Proof. of Proposition 7: We first give the first order condition on A_{EL}^*:

$$
\frac{\partial W_{EL}}{\partial A} = NPV_{EL}\left(\bar{s}_1(A)\right) \tag{73}
$$

$$
+ q_{EL}^* A \frac{\partial \bar{s}_1}{\partial A} \left(1 - \left[\int_{\frac{1}{\alpha}}^1 s f\left(s|\bar{s}_1(A)\right) ds \right] \alpha - \Pr\left[s < \frac{1}{\alpha} | \bar{s}_1(A) \right] \right) g(\bar{s}_1(A))
$$

$$
+ A \frac{\partial q_{EL}^*}{\partial A} \left(\int_\tau^{\bar{s}_1} g(s_1) ds_1 + \int_{\bar{s}_1}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 s f\left(s|s_1\right) ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha} | s_1 \right] \right) g(s_1) ds_1 - \tau\beta \right),
$$

where

$$
NPV_{EL}\left(\bar{s}_1(A)\right) = \tau\beta - 1 + q_{EL}^* \left(\int_\tau^{\bar{s}_1(A)} (1 - \tau\beta) g(s_1) ds_1 + \int_{\bar{s}_1(A)}^1 \left(\left[\int_{\frac{1}{\alpha}}^1 s f\left(s|s_1\right) ds \right] \alpha \right. \right.
$$

$$
\left. \left. + \Pr\left[s < \frac{1}{\alpha} | s_1 \right] - \tau\beta \right) g(s_1) ds_1 \right). \tag{74}
$$

Next we derive the sufficient conditions in Proposition 7. We first prove two preliminary

results that we will use in our future steps:

Result 1: For $A \leq A_{\min}$, $W_{IL} = W_{EL}$. In particular, $\lim_{A \to A_{\min}^-} \frac{\partial W_{IL}}{\partial A} = \lim_{A \to A_{\min}^-} \frac{\partial W_{EL}}{\partial A}$.

Result 1 follows directly from (67).

Result 2: $\lim_{A \to A_{\min}^-} \frac{\partial W_{IL}}{\partial A} > \lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A}$ and $\lim_{A \to A_{\min}^-} \frac{\partial W_{EL}}{\partial A} \geq \lim_{A \to A_{\min}^+} \frac{\partial W_{EL}}{\partial A}$.

The first part of Result 2 holds because from (37),

$$
\lim_{A \to A_{\min}^-} \frac{\partial W_{IL}}{\partial A} - \lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A} \tag{75}
$$

$$
= q_{IL}^* A \frac{\partial \bar{s}_1}{\partial A} \left(\left[\int_{\frac{1}{\alpha}}^1 s f\left(s|\bar{s}_1(A)\right) ds \right] \alpha + \Pr\left[s < \frac{1}{\alpha} | \bar{s}_1(A) \right] - \tau\beta \right) g(\bar{s}_1(A)) > 0.
$$

The second part of Result 2 can be proved analogously. From (73),

$$
\lim_{A \to A_{\min}^-} \frac{\partial W_{EL}}{\partial A} - \lim_{A \to A_{\min}^+} \frac{\partial W_{EL}}{\partial A} \tag{76}
$$

$$
= -q_{EL}^* A \frac{\partial \bar{s}_1}{\partial A} \left(1 - \left[\int_{\frac{1}{\alpha}}^1 s f\left(s | \bar{s}_1(A)\right) ds \right] \alpha - \Pr\left[s < \frac{1}{\alpha} | \bar{s}_1(A) \right] \right) g(\bar{s}_1(A)) \geq 0.
$$

Given Result 1 and Result 2, we derive some sufficient conditions for $A_{EL}^* > A_{IL}^*$ and $W_{EL}(A_{EL}^*) > W_{IL}(A_{IL}^*)$. We proceed in three steps. In step 1, we prove that $\lim_{A \to A_{\min}^+} \frac{\partial W_{EL}}{\partial A} > 0$ and $\lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A} < 0$ is a sufficient condition for $A_{EL}^* > A_{IL}^*$ and $W_{EL}(A_{EL}^*) > W_{IL}(A_{IL}^*)$. In step 2 and 3, we reduce the set of conditions of $\lim_{A \to A_{\min}^+} \frac{\partial W_{EL}}{\partial A} > 0$ and $\lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A} < 0$ into a set of conditions on β, $\Pr\left(s < \frac{1}{\alpha} | s_1 = \tau\right)$ and $g(\tau)$.

Step 1: We prove that $\lim_{A \to A_{\min}^+} \frac{\partial W_{EL}}{\partial A} > 0$ and $\lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A} < 0$ is a sufficient condition for $A_{EL}^* > A_{IL}^*$ and $W_{EL}(A_{EL}^*) > W_{IL}(A_{IL}^*)$. By Assumption 5, $\lim_{A \to A_{\min}^+} \frac{\partial W_{EL}}{\partial A} > 0$ implies that for $A \leq A_{\min}$, $\frac{\partial W_{EL}}{\partial A} > 0$. Therefore, $A_{EL}^* > A_{\min}$. In addition, $\lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A} < 0$ implies that for $A > A_{\min}$, $\frac{\partial W_{IL}}{\partial A} < 0$. Furthermore,

$$
\frac{\partial W_{IL}}{\partial A}\Big|_{A=A_{\min}^-} = \frac{\partial W_{EL}}{\partial A}\Big|_{A=A_{\min}^-} > \frac{\partial W_{EL}}{\partial A}\Big|_{A=A_{\min}^+} > 0. \tag{77}
$$

The first equality uses Result 1 and the second inequality uses Result 2. By Assumption 5, for $A < A_{\min}$, $\frac{\partial W_{IL}}{\partial A} > 0$. Therefore, $A_{IL}^* = A_{\min}$, which implies that $A_{EL}^* > A_{\min} = A_{IL}^*$ and $W_{EL}(A_{EL}^*) > W_{EL}(A_{\min}) = W_{IL}(A_{\min}) = W_{IL}(A_{IL}^*)$. The second equality uses Result 1.

Step 2: We show that, under $\Pr\left(s < \frac{1}{\alpha} | s_1 = \tau\right) = 1$, $\lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A} < 0$ if $g(\tau)$ is

sufficiently large. From (37),

$$
\lim_{A \to A_{\min}^+} \frac{\partial W_{IL}}{\partial A}
$$

$$
= NPV_{IL}(\tau) - q_{IL}^* A_{\min} \frac{\partial \bar{s}_1}{\partial A} \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1 = \tau) \, ds \right] \alpha + \Pr \left[s < \frac{1}{\alpha} \middle| s_1 = \tau \right] - \tau\beta \right) g(\tau)
$$

$$
+ A_{\min} \frac{\partial q_{IL}^*}{\partial A} \left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1) \, ds \right] \alpha + \Pr \left[s < \frac{1}{\alpha} \middle| s_1 \right] - \tau\beta \right) g(s_1) ds_1 \right)
$$

$$
< NPV_{IL}(\tau) - q_{IL}^* A_{\min} \frac{\partial \bar{s}_1}{\partial A} \left(1 - \tau\beta \right) g(\tau)
$$

$$
= \tau\beta - 1
$$

$$
+ q_{IL}^* \left[\left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1) \, ds \right] \alpha + \Pr \left[s < \frac{1}{\alpha} \middle| s_1 \right] \right) g(s_1) ds_1 - \tau\beta \right) - A_{\min} \frac{\partial \bar{s}_1}{\partial A} (1 - \tau\beta) g(\tau) \right]
$$

$$
< q_{IL}^* \left[\left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1) \, ds \right] \alpha + \Pr \left[s < \frac{1}{\alpha} \middle| s_1 \right] \right) g(s_1) ds_1 - \tau\beta \right) - A_{\min} \frac{\partial \bar{s}_1}{\partial A} (1 - \tau\beta) g(\tau) \right]
$$

$$
= q_{IL}^* \left[\alpha \left(\int_{\frac{1}{\alpha}}^1 s dH(s) \right) + H \left(\frac{1}{\alpha} \right) - \tau\beta - A_{\min} \frac{\partial \bar{s}_1}{\partial A} (1 - \tau\beta) g(\tau) \right].
$$

The first equality uses $\bar{s}_1(A_{\min}) = \tau$. The second step uses $\Pr\left(s < \frac{1}{\alpha} \middle| s_1 = \tau\right) = 1$ and $\frac{\partial q_{IL}^*}{\partial A} < 0$ from (56). The third step uses (38). The fourth step uses $\tau\beta < 1$. The fifth equality uses the law of iterated expectations and recall that $H(s)$ is the CDF of s. Note that the term in the fifth equality is negative if

$$
g(\tau) > \frac{\alpha \left(\int_{\frac{1}{\alpha}}^1 s dH(s) \right) + H\left(\frac{1}{\alpha}\right) - \tau\beta}{A_{\min} \frac{\partial \bar{s}_1}{\partial A} (1 - \tau\beta)}
$$

$$
= \frac{\alpha \left(\int_{\frac{1}{\alpha}}^1 s dH(s) \right) + H\left(\frac{1}{\alpha}\right) - \tau\beta}{\dfrac{(1-\tau)(1-\tau\beta)}{\left[\int_{\frac{1}{\alpha}}^1 s \frac{\partial f(s|s_1)}{\partial s_1} \big|_{s_1 = \tau} ds \right] [\alpha - 1 + (\beta - \alpha)\tau] + \left[\int_\tau^{\frac{1}{\alpha}} \frac{\partial f(s|s_1)}{\partial s_1} \big|_{s_1 = \tau} ds \right] (\beta - 1)\tau}}. \tag{78}
$$

59

The second equality uses at $\Pr\left(s < \frac{1}{\alpha}|s_1 = \tau\right) = 1$, $A_{\min} = \frac{1-\tau}{(\beta-1)\tau}E$ and

$$
\begin{aligned}
\frac{\partial \bar{s}_1}{\partial A}\Big|_{A=A_{\min}} &= \frac{1}{E}\frac{\tau(\beta-1) - \left[\int_{\frac{1}{\alpha}}^1 sf\left(s|\tau\right)ds\right](\alpha-1)}{\left[\int_{\frac{1}{\alpha}}^1 s\frac{\partial f(s|s_1)}{\partial s_1}\big|_{s_1=\tau}ds\right]\frac{\alpha-1+(\beta-\alpha)\tau}{(\beta-1)\tau} + \left[\int_{\tau}^{\frac{1}{\alpha}}\frac{\partial f(s|s_1)}{\partial s_1}\big|_{s_1=\tau}ds\right]} \\
&= \frac{1}{E}\frac{\tau(\beta-1)}{\left[\int_{\frac{1}{\alpha}}^1 s\frac{\partial f(s|s_1)}{\partial s_1}\big|_{s_1=\tau}ds\right]\frac{\alpha-1+(\beta-\alpha)\tau}{(\beta-1)\tau} + \left[\int_{\tau}^{\frac{1}{\alpha}}\frac{\partial f(s|s_1)}{\partial s_1}\big|_{s_1=\tau}ds\right]}.
\end{aligned}
\tag{79}
$$

The second equality uses $\Pr\left(s < \frac{1}{\alpha}|s_1 = \tau\right) = 1$ such that $\int_{\frac{1}{\alpha}}^1 sf\left(s|\tau\right)ds = 0$. Note that the

right hand side of (78) is independent of $g\left(\tau\right)$. Therefore, there exists a threshold $\underline{g}$ such

that, under $\Pr\left(s < \frac{1}{\alpha}|s_1 = \tau\right) = 1$, $\lim_{A\to A_{\min}^+}\frac{\partial W_{IL}}{\partial A} < 0$ if $g\left(\tau\right) > \underline{g}$.

Step 3: We show that, under $\Pr\left(s < \frac{1}{\alpha}|s_1 = \tau\right) = 1$, $\lim_{A\to A_{\min}^+}\frac{\partial W_{EL}}{\partial A} > 0$ if β is suffi-

ciently small. From (73),

$$
\begin{aligned}
\lim_{A\to A_{\min}^+}\frac{\partial W_{EL}}{\partial A} &= NPV_{EL}\left(\tau\right) + q_{EL}^* A_{\min}\frac{\partial \bar{s}_1}{\partial A}\left(1 - \left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1 = \tau\right)ds\right]\alpha - \Pr\left[s < \frac{1}{\alpha}|s_1 = \tau\right]\right)g(\tau) \\
&\quad + A_{\min}\frac{\partial q_{EL}^*}{\partial A}\left(\int_{\tau}^1\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right) \\
&= NPV_{EL}\left(\tau\right) + A_{\min}\frac{\partial q_{EL}^*}{\partial A}\left(\int_{\tau}^1\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right) \\
&= \tau\beta - 1 + q_{EL}^*\left(\int_{\tau}^1\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right) \\
&\quad + \frac{A_{\min}}{C''(q_{EL}^*)}\int_{\tau}^1\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right](\alpha-1) - \tau\left(\beta-1\right)\right)g(s_1)ds_1 \\
&\quad \times\left(\int_{\tau}^1\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right) \\
&\propto \tau\beta - 1 + q_{EL}^*\left(\int_{\tau}^1\left(\left[\int_{\frac{1}{\alpha}}^1 sf\left(s|s_1\right)ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right) > 0.
\end{aligned}
$$

The first equality uses $\bar{s}_1\left(A_{\min}\right) = \tau$. The second equality uses $\Pr\left(s < \frac{1}{\alpha}|s_1 = \tau\right) = 1$. The

third equality uses $\bar{s}_1(A_{\min}) = \tau$, (74), $q^*_{EL}(A_{\min}) = q^*_{IL}(A_{\min})$, and (56). The fourth step uses that if β is sufficiently small, $\tau(\beta - 1)$ approaches $\int_\tau^1 \left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right](\alpha - 1)g(s_1)ds_1$ and the term

$$\frac{A_{\min}}{C''(q^*_{EL})} \int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right](\alpha - 1) - \tau(\beta - 1)\right)g(s_1)ds_1$$
$$\times \left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right),$$

approaches 0. By continuity, there exists a threshold $\underline{\beta}$ such that if $\beta < \underline{\beta}$, the sign of $\lim_{A \to A^+_{\min}} \frac{\partial W_{EL}}{\partial A}$ is determined by the sign of the expression

$$\tau\beta - 1 + q^*_{EL}\left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right).$$

The last inequality holds because, by Assumption 6, $W_{EL} > 0$ for $A \leq A_{\min}$. $W_{EL} > 0$ in turn implies that

$$\frac{W_{EL}}{A} = \tau\beta - 1 + q^*_{EL}\left(\int_\tau^1 \left(\left[\int_{\frac{1}{\alpha}}^1 sf(s|s_1)\,ds\right]\alpha + \Pr\left[s < \frac{1}{\alpha}|s_1\right]\right)g(s_1)ds_1 - \tau\beta\right) > 0.$$

In sum, if $\Pr\left(s < \frac{1}{\alpha}|s_1 = \tau\right) = 1$, $g(\tau) > \underline{g}$ and $\beta < \underline{\beta}$, $A^*_{EL} > A^*_{IL}$ and $W_{EL}(A^*_{EL}) > W_{IL}(A^*_{IL})$. By continuity, there exists a set of thresholds $\{\underline{\beta}, \underline{s}_1, \underline{p}, \underline{g}\}$, such that, if $\beta < \underline{\beta}$, $\Pr\left(s < \frac{1}{\alpha}|s_1 \leq \underline{s}_1\right) > \underline{p}$, and $\inf_{s_1 \leq \underline{s}_1} g(s_1) > \underline{g}$, $A^*_{IL} < A^*_{EL}$ and $W_{EL}(A^*_{EL}) > W_{IL}(A^*_{IL})$. Numerical analysis confirms that the set of parameters that satisfy these conditions is non-empty. ∎