Mark-to-Market, Loan Retention, and Loan Origination

Alexander Bleck    Pingyang Gao
UBC            Chicago Booth

Abstract

We study the effects of mark-to-market accounting (MTM) for banks following the originate-to-distribute lending model. Banks have expertise in originating loans but it is costly for them to retain the loans on their books. We study how the accounting measurement of the retained loans affects the banks’ origination and retention decisions. We show that, relative to historic cost accounting (HC), MTM has three consequences. First, it improves the accuracy of loan valuation ex post. Second, it forces banks to retain more risk exposure on their own books. Finally, it can reduce ex-ante origination efforts and lower the average quality of loans in the economy. To the extent that lower loan quality and banks’ excessive risk exposure are two important ingredients for the recent financial crisis, we identify one mechanism through which MTM could contribute to financial crises.

Keywords: Mark-to-market accounting, historic cost accounting, loan quality, financial crisis.

JEL codes: G01, G21, G30, M41.

*We thank Anne Beatty, Sel Becker, Jeremy Bertomeu, Qi Chen, Doug Diamond, Richard Frankel, Andrei Kovrijnykh, Eva Labro, Christian Laux, Christian Leuz, Brian Mittendorf, Raghu Rajan, Korok Ray, Stephen Ryan, Haresh Sapra, Katherine Schipper, Cathy Schrand, Amit Seru, Doug Skinner, Phil Stocken, Lars Stole and participants at the NBER-Sloan Project on Market Institutions and Financial Market Risk, the Carnegie Mellon Accounting Theory Conference, the SUERF and Bank of Spain Conference on Disclosure and Market Discipline, the EIASM Workshop on Accounting and Economics, and workshops at the Bank for International Settlements, the Federal Reserve Banks of Chicago and Philadelphia, UBC, UNC, Wharton, Washington University, Duke University, Chicago Booth, Tilburg, CUHK, UST and Ohio State University for valuable discussions and acknowledge financial support from Chicago Booth and UBC.


1 Introduction

Mark-to-market accounting (hereafter MTM) is the practice of marking the book value of an asset to market prices of the same or comparable assets. In contrast, historical cost accounting (hereafter HC) records the value of an asset at its prior book value (possibly with deterministic adjustments). The economic consequences of accounting measurement rules have long been debated in the literature (e.g., Paton and Littleton (1940); Edwards and Bell (1965); Chambers (1966); Ijiri (1975, 1980); Lim and Sunder (1991)). Since the recent financial crisis, the debate among policy-makers and academics has shifted towards the role MTM plays specifically in the banking industry (e.g., Laux and Leuz (2009, 2010); Bushman (2014, Forthcoming)).¹ In this paper we study the economic consequences of measuring banks’ loans under different measurement rules.

We emphasize the point that the reason why a bank holds a loan on its balance sheet is relevant for understanding the economic consequences of different measurement rules of the loan. In a benchmark case, in which the bank’s loan retention decision is treated as given, we show that MTM increases the loan’s valuation accuracy, which in turn improves the bank’s ex-ante decisions, consistent with the conventional wisdom about MTM. The key in our model is to endogenize the bank’s decision to hold a loan on its balance sheet. To this end, we adopt the well-established “skin in the game” model (e.g., Gorton and Pennacchi (1995); DeMarzo and Duffie (1999)). In the model banks follow the originate-to-distribute lending model (hereafter OTD model): they have expertise in originating loans but seek to sell their loans after origination to avoid the retention’s regulatory cost. In distributing (selling) their loans, however, banks face a “lemons” market problem due to their natural information advantage from origination activities. The good bank (the bank with a good loan) overcomes this friction by keeping “skin in the game”: the good bank retains a portion of its loan to convince the market of its high quality. Since the expected net cash flow of the

loan is higher for the good bank, retention is less costly for the good bank making separation possible. In the resulting separating equilibrium, the bad bank sells its entire loan at a low price, while the good bank endogenously retains a portion of the loan. This baseline model is widely used in the theoretical banking literature and validated empirically (see subsection 5.4).

The endogenous loan retention decisions change the evaluation of accounting measurement rules. The good bank retains a critical fraction of the loan to deter the bad bank from mimicking. As we switch from HC to MTM, any bank that retains the critical portion of the loan can mark its retention to the market price. This implies that the bad bank would receive this revaluation benefit if it were to deviate by holding the critical portion of the loan. In other words, switching from HC to MTM increases the bad bank’s payoff from retention off the equilibrium path. Consequently, the good bank raises its level of costly retention to restore the separating equilibrium.

Compared with HC, MTM creates a trade-off: it enhances the ex-post valuation accuracy of retention but raises the costly equilibrium retention. In particular, we demonstrate three consequences of MTM. First, MTM improves the valuation accuracy of the retained loan, fixing the retained loan. Second, the higher level of retention under MTM (relative to HC) means that banks retain more exposure to the risk of the loans they originate. Finally, MTM can reduce the value of originating good loans, resulting in banks’ lower ex-ante incentive to originate good loans. Banks’ ex-ante origination incentive is increasing in the ex-post valuation accuracy of retention but decreasing in the level of the equilibrium retention. When the regulatory cost is sufficiently high, the latter effect dominates the former and MTM impedes overall ex-ante incentives to generate good loans.

Our paper makes two contributions. First, it contributes to the MTM literature. The appeal of MTM is a straightforward corollary of the triumph of the efficient market hypothesis: market prices reflect all relevant information and thus serve as the best measure of asset
values. By exploiting the price informativeness, MTM improves valuation accuracy and leads to better decisions. We show that MTM not only exploits but also affects the price informativeness. Taking to account this endogenous nature of the price informativeness, MTM could be less efficient than HC.

Our approach responds to Demski’s calls for providing micro-foundations of equilibrium expectations in evaluating accounting policies (Demski (2004)). He argues that “the FASB has [...] a penchant for focusing on a type of transaction and then determining the proper accounting treatment of that transaction. It does not [...] overly concern itself with the supply of transactions if it proscribes one particular accounting treatment. [...] The FASB’s Conceptual Framework strikes me as [...] being built upon a foundation that sidesteps micro foundations of the underlying choices, and largely inadequate for scholarly purpose.”

In our model, the price informativeness is sustained by banks’ retention decisions. Treating the retention as given, MTM unambiguously improves efficiency by enhancing valuation accuracy. However, the banks’ retention decisions are endogenous to accounting rules. Switching to MTM to exploit the price informativeness affects the very retention decision (“the supply of the loan transaction”) that sustains the price informativeness. Thus, MTM improves the ex-post valuation accuracy but at the same time increases the cost of sustaining the price informativeness, resulting in a non-trivial trade-off.

Opponents to MTM often accept the conceptual merit of MTM of providing more information, but challenge its practical implementability. They point out that market prices are often noisy or biased in reflecting the asset’s fundamental values in the presence of market illiquidity, and/or that more information may not be efficient in certain settings (see footnote 3). Neither channel arises in our model by design. First, the market price of the sold loan perfectly reveals the quality of the retained loan in our model. Second, information about the retained loan (valuation accuracy) always improves ex-ante incentives in our model.

---

2 The astute reader will recognize that this discussion is reminiscent of the Lucas Critique that policies designed to exploit the observed empirical relation could change the underlying relation. Demski (2004) explicitly draws the connection to the Lucas Critique.
Yet, we show that MTM could be less efficient than HC due to the endogenous nature of price informativeness. We thus complement the MTM literature by identifying a novel problem of MTM.

There have also been some recent efforts to endogenize the information in prices. Reis and Stocken (2007) are among the earliest efforts towards this direction. They endogenize the pricing of inventory by examining a duopoly setting with production. They show that it is difficult to implement fair value measurements because they are endogenous to the strategic interactions between firms. Bleck and Liu (2007) study a firm’s asset sale decision and show that the accounting rule distorts the decision and thus the rule’s information content. Marinovic (Forthcoming) studies an auction model in which the valuation of an asset on the acquirer’s books affects the acquirer’s bidding strategies. He compares the bidding outcomes under three accounting valuation rules, the purchase method, the exit value method, and the perceived value-in-use method. Plantin and Tirole (2015) consider an agency model in which two performance measures, the public market prices of a similar asset and the asset’s actual sale price, are available. They endogenize market prices from a bidding process and show an interesting externality of using market prices. However, the inefficiency of MTM in their paper comes from the imperfection (noise) of the market prices as a measure of the asset value.

The second contribution of our paper is to provide an accounting perspective on factors that led to the 2007-2008 financial crisis. Our model suggests that MTM could lead to

---

3 Most models of MTM do not endogenize price informativeness. Cifuentes, Shin, and Ferrucci (2005); Allen and Carletti (2008); Plantin, Shin, and Sapra (2008) and Shin (2008) fall into this category, as they assume exogenous illiquidity in the asset market. Burkhardt and Strausz (2009) and Acharya and Ryan (2016) show that MTM inefficiently increases the firm’s risk-shifting incentive (see also Heaton, Lucas, and McDonald (2010)). Caskey and Hughes (2012) show that the use of MTM reduces the risk-shifting incentive and improves investment decisions. Lu, Sapra, and Subramanian (2012) demonstrate that MTM induces a tension between the risk-shifting and the debt-overhang problem with investments. Lin and Lu (2014) also find that MTM induces a tension in the socially optimal investment decision. Corona, Nan, and Zhang (2013) study banks’ voluntary choices of MTM, present an interesting possibility of multiple equilibria, and link the multiplicity to the recent financial crisis. In Otto and Volpin (Forthcoming), MTM reduces the use of other relevant information, while in Gigler, Kanodia, Sapra, and Venugopalan (2013) MTM leads to an overuse of information, with negative consequences for investment.
both deteriorating loan quality and the excessive risk retention on the banks’ books, two prominent ingredients into the development of the financial crisis. The prior two decades in the build-up to the financial crisis witnessed a dramatic increase in the use of MTM for banks, in particular in the securitization business. We provide a brief overview of this development in the Appendix. The expanded use of MTM coincided with the remarkable rise in (low-quality) subprime mortgages and the accumulation of risky retained interests on banks’ balance sheets. Our model shows that MTM could interact with the OTD model and exacerbate the contribution of the OTD model to the financial crisis.

The prior accounting literature on the connection between MTM and the recent financial crisis has mostly focused on contagion effects during the financial crisis (e.g., Allen and Carletti (2008); Plantin, Shin, and Sapra (2008)). For example, Plantin, Shin, and Sapra (2008) assumes the presence of fire sales and then shows that MTM amplifies the fire sale problem by creating a complementarity among banks’ decisions to sell illiquid assets. By forcing banks that did not sell to value their assets at the depressed market price, MTM incentivizes banks to sell earlier for a better price. Thus, asset sales caused by an initial shock affect the market price and MTM feeds the initial price impact back into asset sales, resulting in a loop that amplifies the initial shock in the market. Allen and Carletti (2008) makes a similar point that MTM could cause contagion between two inter-linked industries in the spirit of Allen and Gale (2000). However, Laux and Leuz (2009) and Laux and Leuz (2010) question the empirical validity of this contagion channel.

The rest of the paper is organized as follows. Section 2 describes the model, section 3 presents the equilibria, section 4 states our main results, section 5 considers various extensions to the basic model, and section 6 concludes.
2 Model

2.1 The overview

The model starts with a standard “skin in the game” component whereby the bank endogenously holds some risky loans on its balance sheet. We then augment this model with accounting measurement of the retained loan, and compare the retention and origination decisions under two different accounting measurement regimes.

The timing of the model is as follows. There are three dates: \( t = 0, 1, 2 \). At \( t = 0 \), the bank originates a loan (loan origination). At \( t = 1 \), the bank learns about the loan quality, chooses what fraction of the loan to retain and sells the rest to the loan market (loan distribution). At \( t = 2 \), the loan pays off. The risk free rate is zero and all parties are risk neutral.

There is a continuum of ex-ante identical commercial banks. The representative bank finances risky loans with insured deposits and equity. The bank makes decisions to maximize its equity value. The bank’s equity holders have limited liability. If the bank’s cash flow is insufficient to cover the deposits, the depositors are paid off by the deposit insurance fund. It is well-known that deposit insurance induces the bank to retain excessive risk. To counter this asset substitution incentive, the bank is subject to a capital ratio requirement \( \gamma \) and pays a risk assessment \( c \) for each unit of risky asset it retains. Each $1 risky loan the bank retains on the balance sheet has to be financed at least by \( \gamma \) equity and at most by \( (1 - \gamma) \) deposits. Later we also assume that the risk assessment \( c \) is so large that the bank has no incentive to retain risky loans in the absence of informational frictions. The bank adjusts its deposit and dividend policies to rebalance its capital structure.

In sum, we take the bank’s capital structure arrangement discussed above as given, as is common in the literature (e.g., Allen and Carletti (2008), Lu, Sapra, and Subramanian (2012) and Corona, Nan, and Zhang (2015)). We do this to focus on the bank’s origination
and retention decisions under different accounting measurement regimes. The representative bank follows the OTD model: it has expertise in originating loans but a disadvantage in retaining loans on the balance sheet. We operationalize the OTD business model as follows.

### 2.2 Loan origination

The first component of the OTD model is that the bank has expertise in originating loans. At $t = 0$, the bank originates a risky loan that matures at $t = 2$. The loan principal $B_0$ is recorded as the loan’s book value at $t = 0$. The loan carries an interest rate $\frac{R - B_0}{B_0} > 0$. Thus, the loan’s face value is $R$.\(^4\) For simplicity, we assume that at $t = 2$ the risky loan either pays off $R$ in full or defaults with zero recovery. Denote the loan’s random cash flow by $\tilde{x} \in \{R, 0\}$. The probability of $x = R$ is $\theta \in \{g, b\}$, with $1 > g > b > 0$. We call $\theta$ the quality (type) of the loan, or interchangeably, the quality (type) of the bank. A good loan has a lower default probability than a bad loan.

The loan quality could be improved by the bank’s origination effort at $t = 0$. The bank with origination effort $m$ receives a good loan ($\theta = g$) with probability $m$ and a bad loan ($\theta = b$) with probability $1 - m$, that is $\Pr(\tilde{\theta}(m) = g) = m$. The origination effort $m$ comes at a private cost $s(m)$ to the bank. $s(m)$ satisfies the standard properties: $s(0) = s'(0) = 0$, $s'(m) > 0$ for $m > 0$, $s'(1) = S$, $s'' > 0$, where $S$ is a sufficiently large positive number to ensure that the choice of $m$ is interior. For simplicity we assume that $m$ is observable to the loan market at $t = 1$. With the risk-free rate normalized to zero, we assume that a good loan yields a positive net present value (NPV) while a bad loan generates a negative NPV, that is $gR > B_0 > bR$.

Throughout the paper we refer to a single loan for ease of reference. However, the single loan should be understood as a stand-in for a large portfolio of loans, each with cash flow $\tilde{x} + \tilde{\epsilon}_n$, where $\tilde{x}$ and $\tilde{\epsilon}_n$ are uncorrelated and represent the systematic and the idiosyncratic

\(^4\) The interest rate on the loan satisfies $R = B_0 \cdot \left(1 + \frac{R - B_0}{B_0}\right)$. 

components, respectively. \( \tilde{\varepsilon}_n \) has mean zero. Thus, the aggregate cash flow of the loan portfolio is 
\[
\tilde{x} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} (\tilde{x} + \tilde{\varepsilon}_n).
\]
Moreover, since we refer to a representative bank, the results could be interpreted as the average aggregate results for the economy. For example, the representative bank’s origination effort \( m \) determines the expected quality of the bank’s loan, but it also measures the average quality of loans in the economy by the law of large numbers.

### 2.3 Loan distribution

At \( t = 1 \), the bank privately observes its loan quality \( \theta \) and makes the retention decision.\(^5\) Retention is costly in that the bank pays a risk assessment \( c \) to the regulator for every unit of the risky loan it carries on its books from \( t = 1 \) to \( t = 2 \).\(^6\) This assumption operationalizes the second component of the OTD model that it is costly for the bank to retain risky loans on its books. We discuss the various interpretations of cost \( c \) in section 5, and will stick to the interpretation of \( c \) as a risk assessment by the FDIC for the ease of reference in the rest of the paper.

What complicates the retention decision is the fact that the bank faces the lemons problem in the loan market as a result of its private knowledge of its loan quality \( \theta \), which in turn results from its expertise in loan origination (the first part of the OTD model).

To overcome the lemons problem, the bank adopts a standard “skin in the game” solution. It retains \( k \) homogeneous portion of the loan on its own books and sells the \( 1 - k \) portion to the loan’s market.\(^7\) The loan market responds with a per-unit price \( p(\mathbf{k}) \) for the sold

---

\(^5\) The bank’s information advantage over loan quality is a classic building block in banking theory. Significant delays between the asset generation and the sale, for instance in shelf registrations, explain the information advantage accumulated by the bank (DeMarzo and Duffie (1999)). Specific direct evidence comes from stock market reactions at the time of loan sales in the secondary market (e.g., Berndt and Gupta (2009); Gande and Saunders (2012)). Moreover, the empirical support for “skin in the game” as a solution to the information advantage is indirect evidence, which we provide in subsection 5.4.

\(^6\) We assume away the risk assessment for holding the loan from \( t = 0 \) to \( t = 1 \). Adding such a cost does not affect the results qualitatively because the bank does not learn the loan quality until \( t = 1 \).

\(^7\) An alternative interpretation is the securitization of the loan. The bank places the whole loan in a
portion. As a result, the bank endogenously holds a non-cash asset, i.e., the retained loan, on its balance sheet.

2.4 Accounting measurement

The key issue of interest is how the retained loan is valued on the bank’s balance sheet at $t = 1$. Accounting measurement of the retained loan matters because we assume that the capital ratio requirement $\gamma$ is based on accounting measures of assets, liabilities and capital. This is consistent with practice.\(^8\)

We consider two polar accounting regimes: historical cost (HC) and mark-to-market (MTM). Under HC, the retained loan is recorded at its initial book value $B_0$ (possibly with pre-determined adjustments). Under MTM, the retained loan is marked to the market price of the sold portion of the loan. Denote by $B^A_1$ the per-unit valuation of the retained loan at $t = 1$ under accounting measurement regime $A \in \{H, M\}$. $A = H$ indicates HC and $A = M$ indicates MTM. Since we will compare the equilibria under HC and MTM, all endogenous valuables should be indexed by accounting regime $A$. To ease exposition, we omit the index $A$ unless we need to compare the two regimes.

3 The equilibrium

An equilibrium strategy profile consists of the triplet $\{m^*, k^*, p^*(k)\}$. Since $m$ is observable, the origination decision is subgame perfect. However, since the loan quality $\theta$ is the bank’s private information, we use Perfect Bayesian Equilibrium (PBE) as the solution concept for

---

the retention decision.

3.1 Preliminary: the bank’s equity value

The bank makes decisions to maximize its equity value. Before we solve for the retention decision at $t = 1$, we analyze the bank’s equity value at $t = 1$ and explain how it is affected by accounting measurement. Towards this purpose we treat the bank’s retention decision $k$ and origination decision $m$ as fixed in this subsection. In the next subsections, we examine these two decisions.

The equity value of the bank of type $\theta$ at $t = 1$, denoted by $V_{1\theta}$, is the sum of expected dividends at $t = 1$ and $t = 2$

$$V_{1\theta} = d_1 + E_{1\theta} [d_2 (\bar{x})].$$

(1)

Dividends $d_1$ and $d_2$ are plugs from the bank’s capital structure choices. Given the deposit insurance assumed in our model, the bank always retains the minimum required capital and distributes any excess capital as dividends. Accordingly, the dividends $d_1$ and $d_2$ are such that the capital requirement binds.\(^9\)

<table>
<thead>
<tr>
<th>Row</th>
<th>Events</th>
<th>Cash</th>
<th>Loan</th>
<th>Deposits</th>
<th>Equity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$t = 0$</td>
<td>0</td>
<td>$B_0$</td>
<td>$(1 - \gamma)B_0$</td>
<td>$\gamma B_0$</td>
</tr>
<tr>
<td>2</td>
<td>$t = 1$ after sale before $d_1$</td>
<td>$CASH$</td>
<td>$kB_1$</td>
<td>$(1 - \gamma)kB_1$</td>
<td>$EQUITY$</td>
</tr>
<tr>
<td>3</td>
<td>$t = 1$ after $d_1$</td>
<td>0</td>
<td>$kB_1$</td>
<td>$(1 - \gamma)kB_1$</td>
<td>$\gamma kB_1$</td>
</tr>
<tr>
<td>4</td>
<td>$t = 2$ before $d_2$</td>
<td>$k\bar{x}$</td>
<td>0</td>
<td>$(1 - \gamma)kB_1$</td>
<td>$k [\bar{x} - (1 - \gamma)B_1]$</td>
</tr>
<tr>
<td>5</td>
<td>$t = 2$ after $d_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes: $CASH \triangleq (1 - k) p - kc + (1 - \gamma) (kB_1 - B_0) + e^0$ and $EQUITY \triangleq (1 - k) (p - B_0) + k (B_1 - B_0) - kc + e^0 + \gamma B_0$.

Table 1: The bank’s balance sheets.

We describe the bank’s balance sheets working backward from $t = 2$. After the distribution of dividend $d_1$ at $t = 1$ but before the loan payout, the bank’s balance sheet

\(^9\) The bank’s underlying optimization program and the proof of the optimality of the bank’s dividend decisions are available upon request.
can be read from row 3 in Table 1. The bank holds risky loan with book value $kB_1$ and
finances it with $\gamma kB_1$ equity and $(1 - \gamma) kB_1$ deposits. If the loan pays out $x = R$ per unit,
the bank receives $kR$ from the retained loan. After paying off the deposits $k(1 - \gamma) B_1$, the
bank distributes the remainder as dividend $d_2 = k [R - (1 - \gamma) B_1]$. If the loan defaults and
pays out 0, then the bank equity holders receive $d_2 = 0$ (and the depositors are paid off by
the FDIC). Thus, the dividend $d_2$ is a function of the loan payoff $\bar{x}$

$$d_2 (\bar{x}) = k \max \{ \bar{x} - (1 - \gamma) B_1, 0 \}.$$  (2)

Equation (2) shows that the bank effectively has a call option on the retained loan with
a strike price equal to the deposits $(1 - \gamma) B_1$ (per unit of the retained loan). The loan has
a risky payoff of $\bar{x}$. If the loan pays out $R$, which is greater than the deposits, the bank
exercises the call option by paying off the deposits and pockets the remainder. If the loan
defaults, the bank does not exercise the call option and receives 0. According to the call-put
parity, the call option could be replicated by the underlying loan $\bar{x}$ plus a put option with
the same strike price plus cash, that is

$$E_{1\theta} [\max \{ \bar{x} - (1 - \gamma) B_1, 0 \}] + (1 - \gamma) B_1 = E_{1\theta} [\bar{x}] + E_{1\theta} [\max \{ (1 - \gamma) B_1 - \bar{x}, 0 \}].$$  (3)

It will be clear later that it is more convenient to use the put option formulation.

Now we compute dividend $d_1$. At $t = 1$, the bank starts with a risky loan $B_0$. After
selling $1 - k$ fraction of the loan at price $p$ and marking the retained loan from $B_0$ per unit to
$B_1$ per unit, the bank’s balance sheet is summarized in row 2 in Table 1. The two quantities,
$CASH$ and $EQUITY$, are defined at the bottom of Table 1. With risky loan $kB_1$, the bank
needs $(1 - \gamma)kB_1$ deposits and $\gamma kB_1$ equity, and distributes the excess cash as dividend
$d_1 = CASH$. The amount of $CASH$ results from the following transactions. First, the bank
receives $(1 - k)p$ proceeds from the loan sale. Second, the bank pays risk assessment $kc$.
Third, it raises cash by adjusting the deposit base by $k(1 - \gamma) B_1 - (1 - \gamma) B_0$. Finally, it
receives cash earnings $e^0$ from other sources. We make this assumption so that the bank would not violate the capital requirement at $t = 1$, and thus to avoid the uninteresting case in which the game ends at $t = 1$.

By the double-entry bookkeeping system, dividend $d_1$ can also be computed from the bank’s equity side. The bank’s equity changes as follows. First, the bank recognizes earnings $(1 - k) (p - B_0)$ from the loan sale. Second, the risk assessment cost $kc$ reduces its earnings. Third, the bank recognizes the revaluation of the retained loan, $k (B_1 - B_0)$, if any, in earnings. Finally, it records the earnings $e^0$ from other sources. Dividend $d_1$ is then equal to the beginning balance plus changes minus the ending balance of the equity, that is

$$d_1 = \gamma B_0 + (1 - k) (p - B_0) - kc + k (B_1 - B_0) + e^0 - \gamma k B_1.$$  \hfill (4)

$d_1$ can be verified to be the same as quantity $CASH$ defined in Table 1.

Plugging in $d_1$ and $d_2$ into equation (1) and utilizing the call-put parity from equation (3), the bank’s equity value of type $\theta$ at $t = 1$ can be rewritten as

$$V_{1\theta} (k) = (1 - k) p + k \left( -c + E_{1\theta} [\bar{x}] + E_{1\theta} \left[ \max \{(1 - \gamma) B_1 - \bar{x}, 0\} \right] \right) - (1 - \gamma) B_0 + e^0.$$  \hfill (5)

$V_{1\theta} (k)$ in equation (5) is the key quantity in the model. It describes the bank’s equity value as a function of the retention $k$. For the sold portion of the loan, the bank receives proceeds $(1 - k) p$. For each unit of the retained loan, the bank pays risk assessment $c$, receives the expected cash flow $E_{1\theta} [\bar{x}] = \theta R$, and enjoys a put option $E_{1\theta} \left[ \max \{(1 - \gamma) B_1 - \bar{x}, 0\} \right]$. In addition, the bank receives $e^0$ from other sources and needs to pay off the deposit $(1 - \gamma) B_0$ at $t = 1$.

Note that accounting measurement affects the bank’s equity value $V_{1\theta} (k)$ only through the term $B_1$, the retained loan’s book value at $t = 1$. Moreover, $B_1$ is embedded in the put option value $E_{1\theta} \left[ \max \{(1 - \gamma) B_1 - \bar{x}, 0\} \right]$. Accounting measurement affects the bank’s
equity value by changing the retention’s put option value to the bank’s equity holders. We summarize the properties of the put option value below.

**Lemma 1.** *The put option value is increasing in the book value of the retained loan $B_1$, but decreasing in capital ratio requirement $\gamma$ and loan quality $\theta$.***

These properties are intuitive. The put option arises from the deposit insurance. When the loan valuation is high or the capital requirement is low, the bank is able to finance the same loan with more deposits, raising the strike price of the put option. The bad bank is more likely to exploit the deposit insurance because of its higher default risk.

### 3.2 The retention decision

Having characterized the bank equity value at $t = 1$ in equation (5), we are ready to solve for the bank’s retention $k^*_A$. We focus on separating equilibria whenever possible because loan prices are most informative and thus the motivation for using MTM is the strongest in separating equilibria. In subsection 4.4, we study the pooling equilibria.

From here on, we explicitly index the endogenous variables with accounting regime $A \in \{H, M\}$. $A = H$ indicates HC and $A = M$ means MTM.

In a separating equilibrium, the bank of type $\theta$ retains $k^{A*}_\theta$ and the market perfectly infers the bank’s type $\theta$ from the retention. Denote the market’s belief conditional on observing retention $k$ by $\pi(k) \triangleq \Pr(\theta = g|k)$. Thus, $\pi(k^{A*}_g) = 1$ and $\pi(k^{A*}_b) = 0$. Accordingly, $p(k^{A*}_g) = gR$ and $p(k^{A*}_b) = bR$. It is also convenient to define notations for the put option value when the bank of type $\theta \in \{g, b\}$ holds $k^{A*}_\theta$ under accounting regime $A \in \{H, M\}$, that
The bank’s equity value, defined in equation (5), reveals that there are two components of the equity value that vary with bank type $\theta$. First, the good bank expects to receive $gR$ per unit from the retention while the bad bank expects only $bR$. Second, the put option is more valuable to the bad bank than to the good bank because the bad bank’s loan is more likely to default. Overall, the following lemma shows that the first effect dominates the second. As a result, it is more costly for the bad bank to retain the loan, thus making separation possible.

**Lemma 2.** Retention is more costly for the bad bank than for the good bank.

With these preliminaries, the retention decision can now be characterized as follows.

**Proposition 1.** For $c \geq c_A$, under accounting regime $A \in \{H, M\}$, the unique separating equilibrium is the least-cost one in which

- the retention decisions are $k^{A*}_\theta = \begin{cases} \frac{(g-b)R}{(g-b)R+c-A_b} & \text{if } \theta = g \\ 0 & \text{if } \theta = b \end{cases}$;

- the per-unit loan prices conditional on retention are $p^*(k) = \begin{cases} gR & \text{if } k \geq k^{A*}_g \\ bR & \text{otherwise} \end{cases}$.

Proposition 1 identifies the unique separating equilibrium under each accounting regime $A \in \{H, M\}$. We discuss several features of the equilibrium. First, the separating equilibria arise only when the risk assessment $c$ exceeds the value of the put option the bad bank
receives off equilibrium $A_b$. Otherwise, the retention would always be beneficial to the banks and could not serve as a credible signal of loan quality. The resulting equilibrium would be trivial in that both banks would hold the entire loan.

Second, the bad bank does not retain any loan in equilibrium under either accounting regime, that is $k_b^{A*} = 0$ for $A \in \{H, M\}$. In any separating equilibrium, the bad bank receives $p(k_b^{A*}) = bR$, the worst possible price. Holding any $k_b^{A} > 0$ does not improve the price the bad bank fetches for the sold portion of the loan but incurs a net retention cost $c - A_b$. Thus, the bad bank sells the entire loan in any separating equilibrium. The bad bank’s equilibrium equity value is thus

$$V_{1b}^{A*} \triangleq V_{1b}(k_b^{A*}) = bR - (1 - \gamma) B_0 + e^0, \quad A \in \{H, M\}.$$  \hspace{1cm} (7)

Third, the good bank has to retain enough loan in order to discourage the bad bank from mimicking. To see this, we check the bad bank’s incentive to deviate from $k_b^{A*} = 0$ to $k_g^{A} \geq k_g^{A*}$. The incremental equity value is

$$V_{1b}^{A} (k_g^{A}) - V_{1b}^{A} (0) = (1 - k_g^{A}) (g - b) R - k_g^{A} (c - A_b).$$

Retaining $k_g^{A}$ portion of the loan allows the bad bank to fetch a price premium $(g - b) R$ for the sold portion of the loan, but incurs a net cost of $c - A_b$ for each unit of the retention. As $k_g^{A}$ increases, the benefit becomes smaller (selling a smaller portion) while the cost becomes higher (retaining a larger portion). When $k_g^{A} > k_g^{A*} \triangleq \frac{(g-b)R}{(g-b)R + c - A_b}$, the bad bank does not find it optimal any longer to mimic the good bank. Thus, the bad bank’s off-equilibrium payoff determines the good bank’s equilibrium retention $k_g^{A*}$.

Finally, it can be verified (in the proof) that the good bank has an incentive to retain $k_g^{A*}$, and that the least-cost separating equilibrium $(k_g^{A*}, 0)$ is the unique equilibrium under
Moreover, the good bank’s equilibrium equity value is

\[
V_{1g}^{A^*} \triangleq V_{1g}^{A} \left( k_{g}^{A^*} \right) = (1 - k_{g}^{A^*}) gR + k_{g}^{A^*} (gR - c + A_g) - (1 - \gamma) B_0 + e^0, \quad A \in \{H, M\}. \quad (8)
\]

Having explained the separating equilibrium in Proposition 1, we now discuss its two implications. First, in the separating equilibrium, the loan prices of the sold portion are perfectly informative about the quality of the retained portion of the loan on the banks’ books because they are identical loans. Therefore, there is no illiquidity issue associated with the market price and the implementation of MTM is perfect. Second, the informativeness of the loan prices is not free. It is sustained by the good bank’s costly, inefficient retention.

We now quantify the cost of sustaining the price informativeness. To do this, we need a benchmark in which there is no information asymmetry about loan quality. In this case, \( p(k) = \theta R \) for any \( k \).

**Lemma 3.** For \( c > c_A \), when loan quality \( \theta \) is public information, under both accounting regimes \( A \in \{H, M\} \),

1. both banks sell the entire loan at \( t = 1 \), that is \( k_{g}^{BM^*} = 0 \);
2. the bank’s equilibrium equity value is \( V_{1g}^{BM^*} = \theta R - (1 - \gamma) B_0 + e^0 \);
3. the bank’s ex-ante origination effort \( m_{BM^*} \) is the solution to the first-order condition \( s'(m_{BM^*}) = (g - b)_R \).

In the absence of information asymmetry in the loan market, the bank follows the OTD model. The bank sells the entire loan it originates and there are no risky loans retained on bank balance sheets.

Now we can examine the cost for the banks to sustain the informativeness of the loan price. The bad bank’s equilibrium equity value, computed in equation (7), is the same as that
in the benchmark, that is \( V_{1b}^{BM*} = V_{1b}^{A*} \). This is because in equilibrium the bad bank sells the entire loan at the expected value \( bR \). On the other hand, the good bank’s equilibrium equity values, computed in equation (8), are lower than that in the benchmark. Specifically, the difference is

\[
V_{1g}^{BM*} - V_{1g}^{A*} = k_g^{A*} (c - A_g).
\]

The cost for the good bank to sustain the price informativeness is \( k_g^{A*} (c - A_g) \), the product of the equilibrium retention level and the effective per-unit retention cost. This cost of price discovery is consequential for the bank’s ex-ante origination decision, to which we turn now.

### 3.3 The origination decision

Anticipating that a loan of type \( \theta \) will receive an equity value of \( V_{1\theta}^{A*} \) at \( t = 1 \), the bank chooses origination effort \( m \) at \( t = 0 \) to maximize the expected equity value at \( t = 0 \):

\[
V_A^0 (m) = m V_{1g}^{A*} + (1 - m) V_{1b}^{A*} - s(m).
\]

The optimal origination effort \( m^{A*} \) is the solution to the first-order condition

\[
s'(m^{A*}) = V_{1g}^{A*} - V_{1b}^{A*} = (g - b) R - k_g^{A*} (c - A_g).
\]

(9)

Origination effort \( m \) costs \( s(m) \), but increases the probability of receiving a good loan \( m \). The left-hand side is the marginal cost of origination effort. The right-hand side is the marginal benefit of effort, which is determined by the difference of the equity values of a good and a bad loan at \( t = 1 \). The greater the equity value differential at \( t = 1 \), the stronger the incentive the bank has to originate a good loan at \( t = 0 \). Therefore, the price discovery in the loan market at \( t = 1 \) provides incentives for the loan origination at \( t = 0 \).

The cost of price informativeness reduces \( V_{1g}^{A*} \) and thus lowers the bank’s ex-ante
incentive to originate good loans. Knowing that the payoff of a good loan is lowered by the signaling cost, the bank exerts less effort to improve the loan quality in the first place. To see the last point more clearly, we could compare $m^{A*}$ with $m^{BM*}$

$$s'(m^{BM*}) - s'(m^{A*}) = k_A^*(c - A_g).$$

(10)

The equilibrium origination effort is lower than the benchmark, that is $m^{A*} < m^{BM*}$. Moreover, the shortfall is determined exactly by the signaling cost. MTM and HC affect this signaling cost differently and thus affect origination incentives differently.

4 Analysis

In this section, we compare the economic consequences of MTM relative to HC for banks and the loan market. MTM improves the ex-post valuation accuracy by exploiting the price informativeness. However, in so doing, MTM also makes signaling more costly. As a result, MTM forces banks to retain more risky loans on their own books and could reduce the banks’ ex-ante incentive to originate good loans. In the extreme, we show that MTM could make signaling so prohibitively costly that MTM could destroy the very information in the market price it attempts to exploit.

The subsequent analysis up to subsection 4.4 is based on the separating equilibrium that exists under both HC and MTM for $c > c_M$. The analysis following subsection 4.4 is based on the parameter range $c \in [c_H, c_M]$, in which the separating equilibrium only exists under HC. Figure 1 illustrates.
Figure 1: Range of equilibria

4.1 Benefit of MTM: measurement accuracy

A salient feature of the separating equilibrium under HC in Proposition 1 is that the price of the sold loan perfectly reveals the quality of the retention in equilibrium. Yet under HC, the retention is measured at $B_0$, the loan’s original cost. As such, the retention is systematically under-valued relative to its underlying quality even though the relevant information is available. In contrast, MTM attempts to overcome this measurement deficiency. Under MTM, the retention is measured at $B^*_1 = gR$, the equilibrium (per-unit) price of the sold portion of the loan. Thus, the good bank’s retention is accurately measured.

**Proposition 2.** Under HC, the retention is under-valued in equilibrium. Under MTM, the retention is accurately valued in equilibrium.

Proposition 2 highlights the conventional wisdom of the efficiency of MTM: MTM makes asset valuation more accurate by exploiting the information in asset prices. All else equal, the enhanced measurement accuracy increases the good bank’s equity value, which in turn improves the bank’s ex-ante origination incentive (equation (9)).

4.2 Cost of MTM: risk retention

However, all else is not equal. As we switch from HC to MTM, the equilibrium retention decisions change as well. In particular, the following proposition shows that the equilibrium
retention is higher under MTM than under HC.

**Proposition 3.** Banks retain more loans on their balance sheets under MTM than they do under HC, that is \( k^{M^*}_0 \geq k^{H^*}_0 \) (strict inequality for \( \theta = g \)).

The intuition for Proposition 3 is simple. As the discussion following Proposition 1 suggests, the bad bank’s incentive to mimic the good bank on the off-equilibrium path drives the good bank’s equilibrium retention. When switching from HC to MTM, the only change to the bad bank’s off-equilibrium payoff is that the put option value changes from \( H_b \) to \( M_b \).

By Lemma 1, we know that \( M_b > H_b \) because \( B_1^{M^*} > B_1^{H^*} \), that is, the valuation of the retention (when the bank retains \( k^{M^*}_g \)) is higher under MTM than under HC. By marking the retention to the market price \( B_1^{M^*} = gR \) under MTM, the bad bank could recognize an additional revaluation benefit \( (B_1^{M^*} - B_1^{H^*}) \) for its retention if it were to mimic the good bank by holding \( k^{M^*}_g \). This additional benefit would increase the strike price of the put option and make the bad bank’s deviation more profitable. To deter the bad bank from mimicking, the good bank has to send a stronger, more costly signal by retaining a larger position in the loan.

Proposition 3 demonstrates an important cost of using MTM to exploit the equilibrium informativeness of prices. Recall that the bank would ideally like to dispose of the entire loan. Information asymmetry in the loan market forces the good bank to retain some risky loans in equilibrium. MTM, however, forces the good bank to retain even more than HC. The attempt to exploit the information in price via MTM changes the (signaling) process by which the information is produced and makes the (signaling) process more costly. It is this adverse feedback to the banks’ retention decisions that could reduce the bank’s origination efforts and the overall efficiency of MTM.

Proposition 3 also helps explain the puzzling observation that banks have maintained excessive exposure to the risk of the loans they originated, contrary to what the OTD business model would suggest. This concentration of risk in the banking sector has been alleged as
one of the key factors that turned the subprime mortgage crisis into a full-fledged financial crisis. Banks retain skin in the game to overcome the information asymmetry problem in the loan market. MTM exacerbates the problem by forcing banks to retain more risk exposure on their own balance sheets.

4.3 MTM and loan origination efforts

Having understood how MTM improves the ex-post measurement accuracy but increases the good bank’s costly retention, we turn to its overall effects on the bank’s ex-ante incentive to originate good loans.

**Proposition 4.** There exists a unique threshold of retention cost $\bar{c}$, such that the bank’s origination efforts are lower under MTM than under HC if and only if $c > \bar{c}$.

Proposition 4 gives conditions under which MTM reduces the bank’s ex-ante incentives to originate good loans. The bank’s incentive to originate good loans at $t = 0$ is provided by the difference of the equity value of a good relative to a bad loan at $t = 1$, as we saw in the first-order condition for the origination decision (equation (9)). On the one hand, MTM allows the good bank to mark the retention up to its true value and reduces the effective per-unit retention cost. We call this the unit-cost effect. On the other hand, MTM also forces the good bank to retain more exposure. We term this as the level effect. As a result, the net impact of MTM on the signaling cost is a trade-off between a lower effective per-unit retention cost and a higher retention level.

This trade-off between the unit-cost effect and the level effect is complicated. For example, regulatory cost $c$ both reduces the retention level ($\frac{\partial}{\partial c}k^A_g < 0$) and increases the unit retention cost ($c - A_g$). Proposition 4 shows that when risk assessment $c$ is sufficiently high, the level effect dominates the unit-cost effect. In that case, MTM increases the net signaling cost and reduces the bank’s ex-ante origination effort. The key driving force behind Proposition 4 is the following lemma.
Lemma 4. When switching to MTM, the put option value increases for both the good and the bad bank, but it increases more for the bad bank, that is $M_b - H_b > M_g - H_g > 0$.

The signaling cost is determined by the good bank’s retention level and unit retention cost, but the former is determined by the bad bank’s retention cost. When we switch from HC to MTM, the revaluation of retention reduces the good bank’s unit retention cost on the equilibrium path, but at the same time reduces the bad bank’s unit retention cost off the equilibrium path. Lemma 4 shows that the second effect is more important than the first effect. The reason is that, relative to the good loan, the bad loan is more likely to default at $t = 2$ and thus yields a more valuable put option. To see the intuition better, we could derive

$$m^{M^*} - m^{H^*} \propto k_g^{H^*} (c - H_g) - k_g^{M^*} (c - M_g)$$

$$\propto Y - [(M_b - H_b) - (M_g - H_g)] \cdot c,$$

(11)

where $\propto$ means “has the same sign as”, and $Y$ is a positive constant whose expression is given in the Appendix. From relation (11) and Lemma 4, we know that the equilibrium origination effort under MTM relative to HC is decreasing in $c$. It is then a short step to prove that when the risk assessment is sufficiently high, the equity value of a good loan is smaller, and thus the origination effort is lower, under MTM compared to HC.

Since our bank is a representative bank, the origination effort $m^{A^*}$ also represents the average efforts by all banks in the economy. Thus, it also measures the overall quality of loans in the economy by the law of large numbers. As such, Proposition 4 also provides a link between MTM and the deterioration of the average loan quality in the build-up to the financial crisis.
4.4 MTM and price informativeness

Proposition 3 emphasizes that MTM could make signaling and price discovery more costly. However, as the good bank’s retention increases, it could reach its natural ceiling of \( k = 1 \), the size of the originated loan portfolio. In this case, MTM makes signaling impossible and destroys the informativeness of the price. To study what happens when separation becomes impossible, we make a simplifying assumption that at \( t = 1 \) banks can generate a bad loan at the cost of \( bR \). That is, it is a zero NPV action to generate bad loans ex post. This assumption means that the supply of bad loans is potentially infinite, which helps pin down the price of the pooled loans. Banks will not take advantage of this technology when separation is possible, because in a separating equilibrium a bad loan receives its expected profit and generates zero NPV.

**Proposition 5.** For \( c_M > c \geq c_H \), there does not exist a separating equilibrium under MTM even though there is a unique separating equilibrium under HC. Instead, there exists a pooling equilibrium under MTM in which

- all the good banks retain \( k_g^* = 1 \); the mass of bad banks (per 1 good bank) retaining \( k_b^* = 1 \) is \( q^* = \frac{(1-b)(1-\gamma)gR-c}{c-(1-b)(1-\gamma)bR} \); the rest, if any, retains nothing;

- the per-unit prices of the loans conditional on retention are

\[
p^s(k|q^*) = \begin{cases} 
\frac{2R+q^*bR}{1+q^*} & \text{if } k = 1 \\
 bR & \text{otherwise}
\end{cases}
\]

Proposition 5 complements the separating equilibria under HC and MTM in Proposition 1. We discuss its three features. First, \( c_M > c_H \). It is easier to sustain separation under HC than under MTM. As regulatory cost \( c \) falls, holding retention becomes more attractive. To deter the bad bank from mimicking and to sustain the separating equilibrium, the good bank has to raise its level of retention. When the equilibrium retention reaches its upper bound, there does not exist any pure-strategy separating equilibrium any longer. Because
by Proposition 3, the retention level under MTM $k_g^{M*}$ reaches the upper bound earlier than $k_g^{H*}$, its counterpart under HC. Thus, the separating equilibrium collapses sooner under MTM than under HC.

Second, when separation becomes impossible under MTM, the resulting equilibrium involves pooling: all good banks retain $k_g^{M*} = 1$ and the mass of bad banks retaining $k_b^{M*} = 1$ is $q^*$. In equilibrium, a bad bank is indifferent between retaining 1 and 0, and the market price of the sold loan reflects the average quality of the loan pool $p(k^*|q^*) = \frac{q^*R + q^*bR}{1+q^*}$. Moreover, $p(k^*|q^*)$ is decreasing in $q^*$: the mimicking by bad banks reduces the average quality of the retained loans on the banks’ balance sheets.

Finally, when the separating equilibrium collapses, the loan price in any resulting non-separating equilibrium becomes less informative about the quality of the retained loan. In an attempt to “correct” the inefficiency of HC, MTM exploits the information in the loan price, only to destroy its informativeness in equilibrium. This paradoxical result highlights the endogenous nature of the informativeness of the loan price.

5 Discussion and extensions

We have shown that, in an attempt to exploit price informativeness, MTM makes it more costly to sustain price informativeness and, in the extreme, destroys price informativeness. Another way to see the intuition for this result is as follows. MTM does not reveal the bank’s private information about the quality of the retained loan directly. If MTM could grant the mark-up only to the good bank and force the bad bank to mark its retention down, or equivalently if the information MTM exploits were exogenous, MTM would improve the valuation accuracy without inducing a strategic response. However, the MTM rule stipulates that the retention be marked to the market price of the loan the bank has sold. To exploit the information in the loan price, MTM has to rely on the banks’ retention behavior.
to generate the information. Thus, the information in the loan price MTM exploits is endogenous to the banks’ retention decisions. Because the bad bank could also hold retention on the off-equilibrium path, the early recognition benefit under MTM is also available to the bad bank on the off-equilibrium path. This increase in the payoff for the bad bank on the off-equilibrium path, induced by MTM, increases the equilibrium retention by the good bank and makes signaling more costly. In this sense, the attempt to exploit the information in the loan price via MTM makes producing the information in the loan price more costly.

The basic model illustrates the point that the attempt to exploit the information in asset prices interferes with the market mechanism that sustains the informativeness of price in the first place. It is this feedback effect that could compromise the efficiency of MTM and other market-based policies. In this section, we discuss the robustness of the model to various alternative specifications and provide justifications for some key assumptions.

5.1 Interpretations of retention cost $c$

Cost $c$ plays a crucial role in the model. Conceptually, $c$ reflects the cost excess for the bank, relative to other parties, to hold the loan. In the past three decades, the banking business model has been shifting from the traditional “originate-to-hold” model to the “originate-to-distribute” model (e.g. Bernanke (2008)). This shift is driven by the relative cost of financing loans with internal versus external capital. Berger, Kashyap, and Scalise (1995) “emphasizes regulatory changes and technical and financial innovations as the central driving forces behind transformation of the industry”. Deregulation has increased competition in deposit markets and increased the cost to fund loans with deposits; technical and financial

---

10 By the second quarter of 2008, the outstanding balance of asset-based securities (ABS), including both mortgage and non-mortgage related ABS, is estimated to be $10.24 trillion in the United States and $2.25 trillion in Europe, with an issuance of $3,455 billion in the U.S. and $652 billion in Europe in 2007, according to SIFMA data. Securities Industry and Financial Markets Association (SIFMA), http://www.sifma.org/research/statistics.aspx. In addition, banks also distribute loans through the syndicated loan market and the secondary loan market, which had an annual volume exceeding $1 trillion in the past few years.
innovations reduce the cost to obtain funds from the loan market. As the internal cost of
capital increases and the external cost of capital decreases, it becomes more likely that the
bank that originates the loan is not the best party to hold the loan. We capture this driving
force behind the OTD model by assuming that the bank, relative to investors in the loan
market, incurs an extra cost $c$ for retaining a unit of risky asset on its balance sheet.

One interpretation of the cost $c$ is the regulatory cost imposed on regulated financial
institutions. It could be thought of as the assessment charged by the Federal Deposit
Insurance Corporation (FDIC) in the United States, which is a function of the risk of a
bank’s balance sheet. Alternatively, a typical capital requirement stipulates that banks set
aside a capital reserve for the risky assets on their balance sheets. $c$ reflects the marginal
cost for the bank to meet the capital requirement when they take on one more unit of a risky
asset.

For unregulated financial institutions, the cost $c$ could correspond to any cost differentials
for them and investors to fund loans. For example, $c$ could be interpreted as the cost associated
with the lack of diversification when a financial institution retains all of the loans it originates
on its own books (e.g. Leland and Pyle (1977)). For another example, $c$ could reflect the
relative expertise or investment opportunity of the financial institution and investors in the
loan market. The financial institution has a competitive advantage in originating loans but
other parties (investors) have a competitive advantage in managing the loans; similarly, the
financial institution has other profitable investment projects but faces a financial constraint
while investors have idle capital (e.g. Gorton and Pennacchi (1995)).

5.2 Other alternatives to improve the efficiency of MTM

Given the possible inefficiency of MTM, an interesting question is whether measures to
improve its efficiency exist either in the hands of the regulator or the bank. When $c$ is
interpreted as a regulatory cost, one might naturally wonder if regulators could improve the
efficiency of MTM by linking \( c \) to such observable bank characteristics as the retention level. The optimal design of \( c \) in a general setting is apparently beyond the scope of this paper. Instead, with the interpretation of \( c \) as the FDIC assessment, if we assume that regulators are subject to the same budget constraints under HC and MTM, then, a combination of MTM and any assessment rule that links \( c \) to the retention does not qualitatively change the trade-off of MTM. The intuition is as follows. Indexing \( c \) to the retention is based on the same idea as MTM, namely to exploit the information in loan price. Since regulators have the same information problem investors face, the change in \( c \) cannot be set as a function of the bank’s true type and instead has to be imposed uniformly. Since bad banks benefit more from early recognition, the differential benefit still exists after the regulators increase \( c \) for both banks and indexing \( c \) to retention \( k \) could exacerbate the problem in the same way MTM does.

Banks could take some measures to exploit the information in the retention. For example, banks could resell, hedge, or collateralize the retention after its quality has been established by the signaling game. These issues do not arise in our model because it only spans two periods but would if we allowed for a more elaborate model. However, conceptually these measures share the same idea of exploiting the information in the loan market. We conjecture that the essence of the arguments in the previous paragraph still applies. There is no free information when information has to be sustained by a costly private action.

5.3 Pre-commitment vs. ex-post discretion in loan distribution

In our model banks choose the retention after they learn about the quality of their loans, resulting in ex-post inefficiency of dissipative signaling. An alternative is for banks to use pre-committed retention whereby banks pre-commit to a retention level before they originate loans. An apparent drawback of this pre-commitment is that nothing could prevent the banks from deviating from the commitment after they learn the information ex post. More
subtly, discretionary ex-post retention dominates pre-committed retention when the moral hazard problem in loan origination is severe. The intuition highlights the novel feature of our model that information revealed through ex-post signaling is useful in resolving the moral hazard in the origination effort. Thus, the value of ex-post signaling is greater the more severe the moral hazard problem in the origination. Therefore, the severity of moral hazard in loan origination is an important predictor of the bank’s choice of securitization methods.

5.4 Performance of proportional retention

We model the mechanism of information transmission as proportional loan retention and circumvent the issue of optimal mechanism design, which is apparently beyond the scope of the paper. Common examples of other mechanisms include the design of retained securities (Innes (1990); DeMarzo and Duffie (1999); Fender and Mitchell (2009)) and performance-contingent contracts (Akerlof (1970); Gorton and Pennacchi (1995)). Notwithstanding the difficulty of a comparison of more general mechanisms, we can show any contingent contract to be equivalent to proportional retention provided that the bank’s regulatory cost is proportional to the bank’s risk exposure resulting from the contract. Therefore, our mechanism of proportional retention is not dominated and has the benefits of being simple and particularly well-suited for the accounting issue we study.

A state-contingent contract, such as a performance guarantee, and proportional retention are essentially all examples of signaling/screening. While the labels may differ, the economic forces are similar: they all induce self-selection incentives created by the single-crossing property. That is, since any contractual mechanism tied to the future loan cash flow $\tilde{x}$ necessarily depends on loan type $\theta$ via the cash flow distribution, there will always exist an incentive for the bank of type $g$ to select a contract that is too costly for the $b$ type to also accept thereby revealing their types. The selection of a contract by the $g$ type essentially

\footnote{The proof of this claim is available upon request.}
imposes risk on the bank which is costly for a regulated institution. As long as the bank’s regulatory cost is proportional to the bank’s risk exposure resulting from the contingent contract, any such mechanism is equivalent to proportional retention.

In addition to its simplicity, the key benefit of using proportional retention instead of a general contract is that proportional retention gives a very clean interpretation of MTM. With proportional retention, the portion of the loan that is sold and the portion of the loan that is retained are of identical quality. Therefore, the unit price of the portion of the loan that is sold is the perfect candidate to value the portion of the loan that is retained. This simplifies the model away from the implementation issue of MTM. In contrast, if the bank uses a general contract, e.g. taking the first-loss type of guarantee, the price at which investors buy the loan cannot be directly applied to value the guarantee under MTM. Instead, inference about the quality of loan has to be made from the price of the sold claim and then this inference is used to decide a market value for the guarantee. With a non-proportional contractual guarantee, one has to first undo the distortion created by the shape of the payoff function of the claim to draw an inference about the fundamentals. In other words, the unit price of the sold claim in general differs from the unit market value of the guarantee even though the underlying inferences made about the quality of the two claims are the same. This complexity in implementing MTM is not necessary for the purpose of the paper.

Finally, we use “skin in the game” to motivate the rationale for holding retention and then study the effect of accounting measurement on retention. While it is still debated whether “skin in the game” actually worked as intended, there is both theoretical and empirical support for this assumption. Retaining partial interests by the bank could be a solution to both its information advantage over investors or its unobservable incentive to improve the value of loans (e.g. Leland and Pyle (1977); Gorton and Pennacchi (1995); DeMarzo and Duffie (1999)). There is also empirical evidence indicating that banks do have private information and use retention as a signal (e.g. Simons (1993); Higgins and Mason (2004); Sufi (2007); Keys, Mukherjee, Seru, and Vig (2009); Acharya and Schnabl (2010); Loutskina
and Strahan (2011); Erel, Nadauld, and Stulz (2014); Ashcraft, Gooriah, and Kermani (2014)). Further, retention is typically a large component on a bank’s balance sheet and exerts important influences on a bank’s income statement. Using the data from regulatory filings (e.g. schedules HC-S in Y-9C and RC-S in Call Reports) that U.S. bank holding companies file quarterly with the Federal Reserve, Chen, Liu, and Ryan (2008) report that on average the value of interest-only strips and subordinated asset-backed securities, two components of retention, accounts for about 11% of the outstanding principal balance of private label securitized loans. The information about a bank’s position in retention interest is also available from SEC filings (e.g. 10-Q and 10-K) if the position is material.

5.5 Middle ground: Lower of Cost or Market

Our analysis relies on the comparison of two accounting regimes in their pure forms. In the model, book values under MTM rely solely on current information extracted from market prices while book values under HC do not at all. This choice of pure accounting regimes is intentional to underscore the main theoretical point of the paper. In reality, HC is often implemented using information from market prices in some circumstances in the form of the so-called lower-of-cost-or-market rule (LCM). LCM requires a downward revaluation of the book value of an asset from its current book value but does not allow an upward revaluation. In other words, relative to HC, LCM requires the early recognition of losses (and is thus also known as HC with impairment). Note that in our model LCM would behave in the same manner as HC because the early recognition of losses is not an issue. Rather, the inefficiency in our model under HC manifests itself as the undervaluation of the retention and this undervaluation issue would still exist under LCM.
6 Conclusion

In this paper, we propose a new mechanism by which MTM could impose inefficiency on banks following the OTD model. We show that, relative to HC, MTM could induce banks to retain excessive exposure to the risk of the loans they originated and reduce banks' ex-ante incentive to originate good loans. In the presence of information asymmetry, the informativeness of loan prices is fragile in that it is sustained by a costly signaling process. The attempt to extract information from loan prices makes the signaling process costlier and in the extreme destroys the price informativeness. It is this feedback effect that compromises the efficiency of MTM and causes damage to the real economy. Our paper underscores that information and liquidity in asset markets are not exogenous. Rather, they are determined by the incentives and ability of market participants to overcome market frictions.

The appeal of MTM to economists is a straightforward corollary of the triumph of the efficient market hypothesis: market prices reflect all relevant information and thus serve as the best measure of asset values. By exploiting the information in market prices, MTM values a bank’s assets more accurately. The enhanced valuation accuracy improves the banks’ decisions. Opponents of MTM often accept this conceptual merit of MTM, but challenge its practical implementation. In particular, they point out that the market price is often imperfect in reflecting an asset’s fundamental value in the presence of fire sales or illiquidity (Shleifer and Vishny (1992)), and that the rigid reliance on imperfect prices under MTM could lead to undesirable consequences. Since the loan market is perfectly liquid from investors’ perspective, the implementation difficulty associated with exogenous illiquidity does not arise in our model. Instead, our model highlights a conceptual difficulty with MTM. The attempt to exploit the equilibrium price informativeness compromises the informativeness.

This conceptual difficulty with MTM may be a more general feature of accounting measurement beyond our particular setting of the OTD model. In general, a firm’s business model is viable only if it has some competitive advantage over the market in conducting its
activities. As a result, the core assets and liabilities on a firm’s balance sheet, dictated by its business model, are often subject to the same market frictions that sustain the business model. Market prices in these markets are thus endogenously linked to the firm’s activities that are guided partially by accounting measurement. The optimal design of an accounting measurement rule thus requires an understanding of the firm’s business model to efficiently support the business model. The inefficiency of MTM in the OTD model, where retention serves as a costly signal to overcome the information asymmetry in the market, is that MTM treats the retention as if it was sold, directly contradicting the bank’s purpose (business model) of holding retention. Attempting to resolve accounting measurement problems via a market-based solution could lead to unintended and sometimes undesirable consequences.

References


A Appendix

Accounting treatment for retained interest in securitizations

In general, the shift from HC to MTM has accelerated during the past decade. This Appendix describes the accounting for retained interests of securitizations.
Conditional on sales accounting, FAS 140 stipulates the accounting treatment on the transaction date. The subsequent revaluation depends on how the retained interest, a security, is classified. Securities can be classified as trading, available-for-sale (AFS), or held-to-maturity (HTM), with different accounting treatments (FAS 115 and FAS 157). The only restriction FAS 140 imposes on subsequent classification is that prepayment-sensitive securities be classified as either trading or AFS. For simplicity, we assume that the loan is measured at cost before the transaction.

On the transaction date, items could be classified into two overlapping categories for accounting purposes: proceeds received and retained interest. Proceeds received include cash and any other assets obtained, such as derivatives received that do not use the transferred assets as underlying assets. Liabilities incurred, including recourse commitments, are both proceeds and retained interest. Other retained interests include interests in transferred assets, such as proportional holding, interest-only strips (IO), subordinated securities, and Mortgage Servicing Rights (MSRs).

![Classification of considerations from a securitization](image)

**Figure 2:** Classification of considerations from a securitization

For accounting purposes, retained interests that are not proceeds are recorded at pro-rated cost at inception (the proration is based on fair value). The rationale is that the firm has not relinquished its control over these assets and therefore these are not considered to have been sold yet. However, this rationale is overwritten when the retention is classified as
an AFS or trading security and thus FAS 115 and FAS 157 apply.

The proceeds are fair valued at inception. The firm receives these assets or assumes these liabilities as considerations for the sale. FAS 156 requires the fair value option for MSRs at inception (afterwards firms can choose whether to measure MSRs at impaired cost or fair value) and therefore treats MSRs as proceeds. FAS 166 further requires that all assets obtained and liabilities incurred in a securitization be initially measured at fair value. Thus, for accounting purposes, there are no retained interests that are not proceeds after FAS 166.

Subsequently, the accounting treatment of retained interests as well as the proceeds depends on their classification. FAS 140 does not directly govern the classification; instead, FAS 115 and FAS 157 apply. The only requirement of FAS 140 is that prepayment-sensitive securities could not be classified as HTM. It can only be prepayment sensitive if the underlying loans are subject to prepayment (e.g. residential mortgages but not commercial mortgages). Therefore, not only the retained interests but also the proceeds could be revalued either at impaired cost or at fair value. Most big banks choose fair value. The incurred liabilities could be subject to FAS 5 Loss Contingency.

The transferability of the retained interests is typically not restricted in securitizations. Banks could transfer the retained interests, including selling MSRs or securitizing the IOs. This transferability does not contradict skin in the game. If the retention was previously used for signaling, banks wouldn’t be able to sell it at a price commensurate with “high retention”. As a result of this transferability and the FAS 140’s requirement that prepayment-sensitive retained interests couldn’t be classified as HTM, retained interests are rarely classified as HTM.

**Proof of Lemma 2**

From expression (5), we have \( \frac{\partial}{\partial k} V_{16}^A (k) = -c - p + E_{18} [\bar{x}] + E_{18} [\max \{ (1 - \gamma) B_1^A - \bar{x}, 0 \}] = -c - p + \theta R + (1 - \theta) (1 - \gamma) B_1^A. \) Since \( R > B_1^A \), we have \( \frac{\partial}{\partial k} V_{1g} (k) = \frac{\partial}{\partial k} V_{1b} (k) = \)
With this pricing strategy, the complete set of separating PBE levels of retention is determined by the incentive compatibility constraints of both types. These constraints dictate that neither type must have an incentive to deviate from their equilibrium levels of retention \((k^A_g > 0, k^A_b = 0)\)

\[
V^A_{1b} (0) \geq V^A_{1b} (k^A_g) \quad \Rightarrow \quad k^A_g \geq \frac{p (k^A_g) - p (0)}{c + p (k^A_g) - E_{1b} [\tilde{x}] - A_b},
\]

\[
V^A_{1g} (k^A_g) \geq V^A_{1g} (0) \quad \Rightarrow \quad k^A_g \leq \frac{p (k^A_g) - p (0)}{c + p (k^A_g) - E_{1g} [\tilde{x}] - A_g},
\]

where \(E_{1g} [\tilde{x}] = gR\) and \(E_{1b} [\tilde{x}] = bR\). The resulting set of separating PBE levels of retention under accounting regime \(A\) is

\[
S^A_s \triangleq \left\{ (k^A_g, k^A_b) \mid k^A_b = 0, k^A_g \in \left[ \frac{(g - b) R}{c + (g - b) R - A_b}, \min \left[ 1, \frac{(g - b) R}{c - A_g} \right] \right] \right\}.
\]

\(S^A_s\) is nonempty because, using expression \((6)\), \((g - b) R - A_b + A_g = (g - b) \left[ R - (1 - \gamma) B^A_1 \right] > 0\).

We now apply the Intuitive Criterion to refine away all PBEs that involve \(k^A_g \in \left( \frac{(g - b) R}{c + (g - b) R - A_b}, \min \left[ 1, \frac{(g - b) R}{c - A_g} \right] \right)\). For any such \(k^A_g\), consider the deviation \(\tilde{k}^A = \frac{(g - b) R}{c + (g - b) R - A_b}\). The Intuitive Criterion requires that \(\pi (b | \tilde{k}^A) = 0\) because \(\tilde{k}^A\) is an equilibrium-dominated action for type \(b\) by condition \((12)\) (the most favorable belief of investors upon observing \(\tilde{k}^A\) is \(g\)). This implies that \(\pi (g | \tilde{k}^A) = 1\). Given this belief, \(\tilde{k}^A\) is a profitable deviation for type \(g\) by condition \((13)\). Thus, any \(k^A_g \in \left( \frac{(g - b) R}{c + (g - b) R - A_b}, \min \left[ 1, \frac{(g - b) R}{c - A_g} \right] \right)\) does not survive.
The only PBE that survives is \( k_g^A = \frac{(g-b)R}{c+(g-b)R-M_b}, k_b^A = 0 \). Finally, for \( c \geq c_A \triangleq A_b = (1-b)(1-\gamma)B_1^{A*} \), we have \( k_g^A \in (0,1] \) for \( A \in \{ H, M \} \).

**Proof of Lemma 3**

With full information, \( p(k) = \theta R \) and \( E_{1\theta} \{ \max \{(1-\gamma)B_1 - \tilde{x}, 0\} \} = A_\theta \). From equation (5), a sufficient condition for \( \frac{\partial}{\partial R} V_{1\theta}(k) = -c + (1-\theta)(1-\gamma)B_1 < 0 \) for all \( \theta \in \{ g, b \} \) is \( c > (1-b)(1-\gamma)B_1 = A_b \). Thus, \( k_b^{BM*} = 0 \), and \( V_{1\theta}^{BM*} \) and \( s'(m^{BM*}) \) follow immediately from equations (5) and (9), respectively.

**Proof of Proposition 3**

For \( c \geq c_M \), we have \( k_g^{BM*} - k_g^{H*} = \frac{(g-b)R(\gamma - H_g)}{(c+(g-b)R-H_b)(c+(g-b)R-M_b)} - \frac{(g-b)R(1-b)(1-\gamma)(gR-B_0)}{(c+(g-b)R-H_b)(c+(g-b)R-M_b)} > 0 \), where the last equality follows from equation (6).

**Proof of Lemma 4 and Proposition 4**

From equation (6), \( M_\theta - H_\theta = (1-\theta)(1-\gamma)(gR - B_0) > 0 \). Moreover, \( (M_b - H_b) - (M_g - H_g) = (g-b)(1-\gamma)(gR - B_0) > 0 \). This proves Lemma 4.

The value differential of a good loan under MTM and HC is

\[
\Delta(c) \triangleq V_{1g}^{M*} - V_{1g}^{H*} = k_g^{H*}(c - H_g) - k_g^{M*}(c - M_g) = \frac{(g-b)R[(g-b)R(M_g - H_g) + M_bH_g - M_gH_b - c[(M_b - H_b) - (M_g - H_g)]]}{(c + (g-b)R - M_b)(c + (g-b)R - H_b)} \propto Y - c[(M_b - H_b) - (M_g - H_g)],
\]

where \( Y \triangleq (g-b)R(M_g - H_g) + M_bH_g - M_gH_b \). Using expression (6), we have \( Y = (g-b)R(1-g)(1-\gamma)(gR - B_0) > 0 \). Lemma 4 shows \( (M_b - H_b) - (M_g - H_g) > 0 \). This proves the claim in expression (11).
Using expression (6) in equation (14), we can rewrite $\triangle (c)$ as

$$
\triangle (c) = \frac{(1 - \gamma) (g - b)^2 R (g R - B_0) ((1 - g) R - c)}{(c + (g - b) R - M_b) (c + (g - b) R - H_b)}.
$$

It is clear that $\triangle (c) < 0$ if and only if $c > (1 - g) R$. Combined with the condition of $c > c_M$, under which the separating equilibrium exists in Proposition 1, we have proved that there exists $\bar{c} \triangleq \max \{ (1 - g) R, c_M \}$ such that $\triangle (c) < 0$ if and only if $c > \bar{c}$.

The lower origination effort under MTM compared to HC follows from differencing conditions (9) for $A \in \{ H, M \}$, that is

$$
s' \left( \frac{M^*_M}{M^*_H} \right) - s' \left( \frac{M^*_H}{M^*_M} \right) = V_{1_g} - V_{1_b} = \left( V_{1_g} - V_{1_b} \right) = \triangle (c),
$$

where the last equality follows from $V_{1_g} = V_{1_b}$ (see equation (7)). Combining $s' > 0, s'' > 0$ and $V_{1_g} < V_{1_b}$ for $c > \bar{c}$ gives the desired result.

**Proof of Proposition 5**

We verify that the pooling PBE in the proposition is indeed an equilibrium. We express $p(k|\pi(q^*)) = p(k|q^*)$ to highlight the price determination based on $q^*$. Given the retention strategies and the equilibrium beliefs, that is, all good banks and mass $q^*$ of bad banks retain $k_g^{M^*} = k_b^{M^*} = 1$ and the equilibrium belief is $\pi (g|1) = \frac{1}{1 + q^*}$, the prices reflect the average quality of the loan pool. Therefore, $p(k|g^*) = p \left( \frac{1}{1 + q^*} \right) = \frac{1}{1 + q^*} g R + \left( 1 - \frac{1}{1 + q^*} \right) b R$ if $k = 1$ and $p(k|q^*) = b R$ otherwise. Given these prices, we show that the retention strategies $(k_g^{M^*} = 1, k_b^{M^*} = 1)$ satisfy the IC constraints of both types

$$
V_{1_b}^M (1) \geq V_{1_b}^M (0) \quad (15)
$$

$$
V_{1_g}^M (1) > V_{1_g}^M (0) \quad (16)
$$

Therefore, neither type has an incentive to deviate from their equilibrium strategies. Constraint (15) holds with equality at the equilibrium mass of bad banks $q^*$ that are indifferent between retaining 1 and retaining 0, that is for $q^* = \frac{(1 - b)(1 - \gamma)g R - c}{c - (1 - b)(1 - \gamma)b R}$. Constraint (16) holds
because $R > p(1|q^*)$.

**Variable definitions**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expression</th>
<th>Defined on page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_A$</td>
<td>$(1 - b)(1 - \gamma) B_1^{A^*}$</td>
<td>40</td>
</tr>
<tr>
<td>$c_H$</td>
<td>$(1 - b)(1 - \gamma) B_0$</td>
<td>40</td>
</tr>
<tr>
<td>$c_M$</td>
<td>$(1 - b)(1 - \gamma) gR$</td>
<td>40</td>
</tr>
<tr>
<td>$\bar{c}$</td>
<td>$\max{(1 - g) R, (1 - b)(1 - \gamma) gR}$</td>
<td>41</td>
</tr>
<tr>
<td>$A_\theta$</td>
<td>$(1 - \theta)(1 - \gamma) B_1^{A^*}$</td>
<td>6</td>
</tr>
<tr>
<td>$M_g$</td>
<td>$(1 - g)(1 - \gamma) gR$</td>
<td>6</td>
</tr>
<tr>
<td>$M_b$</td>
<td>$(1 - b)(1 - \gamma) gR$</td>
<td>6</td>
</tr>
<tr>
<td>$H_g$</td>
<td>$(1 - g)(1 - \gamma) B_0$</td>
<td>6</td>
</tr>
<tr>
<td>$H_b$</td>
<td>$(1 - b)(1 - \gamma) B_0$</td>
<td>6</td>
</tr>
</tbody>
</table>