High Bank Leverage, Risk Management, and Liquid-Claim Production

Harry DeAngelo and René M. Stulz*

April 2013
Revised October 20, 2013

Abstract

Liquidity production is a central function of banks. High leverage is optimal for banks in a model that has just enough frictions for banks to have a meaningful role in liquid-claim production. The model has a market premium for (socially valuable) safe/liquid debt, but no taxes or other traditional motives to lever up. Because the liquidity premium is paid only for safe debt, banks in the model use risk management to maximize their capacity to include such debt in their capital structures. The model can explain (i) why banks have higher leverage than most operating firms, (ii) why risk management is central to banks' operating policies, (iii) why bank leverage increased over the last 150 years or so, and (iv) why leverage limits for regulated banks impede their ability to compete with unregulated shadow banks.

Contact information: hdeangelo@marshall.usc.edu and stulz@cob.osu.edu

*Respectively, Kenneth King Stonier Chair at the Marshall School of Business, University of Southern California, and Everett D. Reese Chair of Banking and Monetary Economics, The Ohio State University, and NBER.

We thank Anat Admati, Michael Brennan, Charles Calomiris, John Cochrane, Peter DeMarzo, Doug Diamond, Nicola Gennaioli, Charles Goodhart, Gary Gorton, Martin Hellwig, Paul Pfleiderer, Jim Poterba, Aris Protopapadakis, Andrei Shleifer, Richard Smith, Jeremy Stein, and Anjan Thakor for useful comments. Brian Baugh and Yeejin Jang provided excellent research assistance. This paper was formerly titled "Why High Leverage is Optimal for Banks." René Stulz serves on the board of a bank that is affected by capital requirements and consults and provides expert testimony for financial institutions.

High Bank Leverage, Risk Management, and Liquid-Claim Production

1. Introduction

Banks maintain capital structures with leverage ratios that are much higher than those of virtually all non-financial firms. Many economists accordingly see high bank leverage as puzzling from a positive-theory viewpoint and as normatively troubling. These reactions arise from viewing bank capital structure through the lens of Modigliani and Miller (1958, MM) augmented by consideration of moral hazard, taxes, and other leverage-related distortions. MM's debt-equity neutrality principle states that, absent frictions and holding operating policy fixed, all capital structures yield identical value. When leverage-related distortions are added to the debt-equity neutrality baseline, the resultant capital structure model has no efficiency-based motive that can explain why banks generally maintain leverage ratios that are so much higher than those of operating firms.

Admati and Hellwig (2013), building on Miller (1995), Pfleiderer (2010), and Admati, DeMarzo, Hellwig, and Pfleiderer (2011), use this capital structure model to argue for severe regulatory limits on bank leverage. As Myerson (2013, p. 3) discusses, MM's leverage irrelevance theorem is the foundation of the argument. With debt-equity neutrality as the baseline and only leverage-related distortions given meaningful weight, Admati and Hellwig (2013, p. 191) conclude: "increasing equity requirements from 3 percent to 25 percent of banks' total assets would involve only a reshuffling of financial claims in the economy to create a better and safer financial system. There would be no cost to society whatsoever." Cochrane (2013) endorses this general view of bank capital structure and notes that the argument favors increasing bank equity requirements to 50% or even 100% of assets.[1]

In the MM baseline view, there is no connection between bank leverage and what banks do. Banks are treated as firms that make loans, and they make the same loans irrespective of their debt-equity mix. Importantly, with its central principle that capital structure is irrelevant, this view leaves no room for a

---

[1] This general view has strong support among many other prominent economists. For example, seventeen other well-known economists agree that, with much more equity funding, banks could perform all their socially useful functions and support growth without endangering the financial system. See "Healthy banking system is the goal, not profitable banks," letter published in the Financial Times, November 9, 2010. See also Myerson (2013) and Wolf (2013) for strong endorsements of much larger equity capital requirements for banks.

connection between bank leverage and the value that banks generate as producers of liquid financial claims. The idea that liquidity production is intrinsic to banking is discussed extensively by, among others, Diamond and Dybvig (1983), Diamond and Rajan (2001), Gorton (2010), Gorton and Pennacchi (1990), and Holmström and Tirole (1998, 2011). If banks' credit-screening technology enables them to make better loans than competitors could and all other MM assumptions hold, banks could adopt all-equity capital structures with no loss in value. However, if banks generate value by producing financial claims to meet the demand for liquidity, those with high-equity capital structures are not competitive with otherwise comparable banks or bank substitutes that have less equity.

In this paper, we show that high leverage is optimal in a model of bank capital structure that has just enough frictions so that banks have a socially valuable role in supplying liquid claims (safe debt) to parties with imperfect access to capital markets. To establish that high bank leverage is the natural (distortion-free) result of intermediation focused on liquid-claim production, our model rules out taxes, agency problems, deposit insurance, reaching-for-yield behavior, ROE-based compensation plans, and all other such factors that might encourage banks to include debt in their capital structures.

The foundational elements of this idealized (stripped-to-the-basics) model of intermediation are: (i) inclusion of an exogenous demand for liquid claims in the spirit of Diamond and Dybvig (1983) and the other pioneering studies referenced immediately above, (ii) the existence of costs of intermediation that are a function of bank scale, and (iii) the ability to engage in asset-side risk management strategies that involve hedging in a perfect/complete capital market, exactly as in models that yield the MM theorem.

Together, (i) and (ii) imply that a liquidity premium – a rate-of-return discount – on safe/liquid debt can obtain because scale-related costs of banking preclude the unfettered arbitrage that would make intermediation irrelevant. The existence of a liquidity premium is a common assumption in models of intermediation, and the available evidence supports this approach. In our model, a premium obtains under some, but not all, conditions. For example, if the demand for liquid claims is small relative to the efficient size of the banking sector, liquidity does not command a premium. The leverage of any one bank is then irrelevant, but aggregate bank debt matters and is tailored to match liquid-claim demand. If liquid-claim demand is strong and bank scale is determinate, then liquidity is priced at a premium and high leverage is optimal for individual banks and banks in the aggregate.

Condition (iii) is implicitly present in many prior analyses of bank capital structure. It is a previously unrecognized *asset*-side implication of models that apply the MM theorem to the *liability* side of bank balance sheets. When market prices embed a liquidity premium, banks in our model generate value by exploiting hedging opportunities that are made possible by condition (iii) to construct suitably safe asset portfolios to support safe/liquid debt issued to parties with imperfect access to capital markets.

When a liquidity premium exists, bank capital structure matters: Equity and safe/liquid debt are not equally attractive sources of capital. As Gorton and Pennacchi (1990) emphasize, debt has a strict advantage when it has the informational insensitivity property – immediacy, safety, and ease of valuation – desired by those seeking liquidity. In our model, this property applies only to perfectly safe deposit debt. However, the reasoning of Gorton and Pennacchi implies that liquidity benefits can also be priced into relatively safe debt other than riskless bank deposits per se. Hence the incentive that we discuss for risk management to support banks' ability to lever up with safe debt arguably generalizes to capital structures with more debt than simply bank deposits.

Capital structure is a sideshow for value creation at operating firms, but it is *the* star – actually co-star – of the show at banks in our model. The risky asset structures of most operating firms are poorly suited to support large-scale issuance of safe/liquid debt. In our model, banks exist because specialization and the associated cost efficiencies give them a comparative advantage over operating firms in arranging asset structures to support capital structures with abundant quantities of liquid financial claims.

Risk management – specifically, asset-structure optimization focused on reducing risk – is the other co-star, and it is the critical feature that gives banks a comparative advantage in producing liquid claims. Simply levering up with a risky asset structure does *not* generate greater value as it would, e.g., even for firms with highly risky asset structures in the MM corporate tax model. The reason is that the liquidity premium is paid only for safe debt. Risk management that expands the capacity to issue safe debt is what enables banks to generate value by adopting high leverage capital structures in our model.

The general point is that, given a material demand for liquidity per se, intermediaries will emerge to meet that demand with high-leverage capital structures, which are made possible by asset-allocation choices optimized under conditions of uncertainty for maximal production of safe/liquid claims. This is the fundamental reason why debt-equity neutrality is an inappropriate equity-biased baseline for analyzing bank capital structure. Debt-equity neutrality obtains in the MM model where banks have no reason to exist. When we instead consider a stripped-to-the basics model that gives banks a meaningful role in liquid-claim production, high leverage coupled with low asset-side risk emerges as the baseline optimum expected of banks before one considers taxes and other leverage-related distortions.

The analysis thus cautions against accepting the view that high bank leverage must be the result of moral hazard, other agency problems, or tax motives to borrow. In our model, there are no such motives to lever up, yet high bank leverage is optimal. High leverage is the result of intermediation that is focused on the optimal production of (privately and socially beneficial) liquid financial claims.

The analysis also cautions against concluding that bank leverage must be too high because operating firms maintain much lower leverage. In our model, banks engage in risk management to support greater production of safe/liquid financial claims, which results in high leverage. In contrast, operating firms create value through real project choices, which are not well suited to support large amounts of safe debt because they commonly entail significant cash flow uncertainty.

In addition to providing a simple fundamentals-based explanation why banks have higher leverage ratios than most operating firms, our model also yields some interesting ancillary implications about bank capital structure. For example, greater competition that squeezes bank liquidity and loan spreads diminishes the "franchise" value of (infra-marginal) banks, and the resultant reduction in equity value implies an increase in optimal bank leverage ratios. Also, if conventional banks face regulatory limits on leverage while shadow banks do not, the former will be at a competitive disadvantage to the latter, and liquid-claim production will migrate into the unregulated shadow-banking sector.

An important caveat is that real-world banks face constraints and incentives that can make them unable and/or unwilling to implement the risk-management strategies they adopt under our idealized conditions. Therefore, in capital structure models that move beyond our basic setting, there can be legitimate, and potentially substantial, benefits to regulations that limit bank leverage. Proper evaluation of such regulations requires weighing the costs and benefits associated with the various frictions that affect real-world banks, but are excluded from our model. Our model is therefore silent about whether such regulations are warranted. It simply highlights the possibility that such regulations could impair production of socially valuable liquidity, and perhaps exacerbate systemic risk by inducing a substitution of liquid-claim production into the unregulated shadow-banking sector.

Our paper is not the first to recognize the inapplicability of the MM theorem to bank capital structure decisions when liquidity is priced at a premium. Hanson, Kashyap, and Stein (2011, p. 17 and fn 1) and Flannery (2012) highlight this point, and it plays a major role in important recent papers by Stein (2012) and Gennaioli, Shleifer, and Vishny (2012, 2013). Our paper is differentiated by its focus on banks' incentive to use risk management strategies to expand their capacity to issue large amounts of safe/liquid debt. Optimal risk management – coupled with an MM violation when liquidity per se is valuable – gives banks the incentive to lever up and generate value from the liability side of their balance sheets.

Gorton and Pennacchi (1990) have previously recognized that asset diversification could help foster the production of safe/liquid debt. Gennaioli, Shleifer, and Vishny (2013) also recognize the incentive to diversify to support safe debt issuance, but they focus on a setting in which asset diversification is imperfect and therefore leaves banks exposed to correlated tail risk from related loan and security holdings. Other studies note the benefits of bank diversification outside the context of liquid-claim production. For example, in Diamond's (1984) delegated-monitoring model, greater asset diversification enables banks to issue more debt, which dominates equity when banks generate value by monitoring borrowers on behalf of capital suppliers. More recently, Gornall and Strebulaev's (2013) tax/bankruptcy cost model of capital structure implies that high bank leverage arises from low portfolio volatility due to diversified holdings of senior claims on borrower assets.

Section 2 describes the model's basic elements. Section 3 characterizes optimal risk management for liquid-claim production. Section 4 analyzes optimal bank capital structure holding scale fixed and taking as given a premium for liquidity, while section 5 treats bank scale as endogenous. Section 6 considers bank leverage and systemic risk when risk-management costs make it prohibitive for banks to produce perfectly safe claims. Section 7 concludes.

2. Basic model details

Our model includes just enough frictions for banks to have a meaningful role in liquid-claim production. There are no taxes, agency costs, or other leverage-related distortions. Nor is there deposit insurance. We also exclude informational asymmetries, e.g., as in the analysis of debt, asset collateral, and asset-quality disclosures in the intermediation models of Dang, Gorton, Holmström (2013) and Dang, Gorton, Holmström, and Ordonez (2013).

We assume the co-existence of two sharply segmented financial markets. Market I is populated by operating firms, households, and financial intermediaries that transact under the perfect/complete market conditions that lead to the MM theorem.[2] Market II is populated by a completely distinct set of agents that do not have access to market I and that can trade among themselves, but not under perfect/complete market conditions. Given market II's imperfections, there are firms and households in II that willingly pay a premium (over the relevant price in market I) for liquid financial claims, i.e., claims that provide immediate, assured access to capital.

There is, of course, no reason for banks or other intermediaries to exist in a model that has only market I. Banks could still arise as "neutral mutations" that raise funds in market I and costlessly repackage claims for agents that could, at no cost, do on their own what banks do for them. If only market I were operative, bank capital structure would not matter, while banks themselves would be redundant and generate no benefits.

With both markets I and II in our model, banks serve a useful purpose by bridging the gap between the two markets and supplying debt (deposits) to the parties in II that are willing to pay a premium for

[2] Complete markets are sufficient but not necessary for the MM theorem. Our arguments go through unchanged if the bank has access to incomplete but otherwise frictionless markets in which a riskless claim can be constructed from the set of existing securities in market segment I. See section 3.

safe/liquid claims. Specifically, banks can issue safe debt to agents in market II that value having assured access to capital, while backing that safe debt with assets acquired at lower cost in market I using resources obtained through debt issuance.

When a liquidity premium is earned by supplying safe debt to market II, operating firms with access to market I have incentives to compete with banks in supplying such debt. However, their ability to compete is generally limited because the earnings from real production are inherently risky and therefore unable to support capital structures with large amounts of safe debt. Moreover, we assume that banks have a comparative advantage over operating firms both in (i) managing asset-side risk to foster production of safe debt and in (ii) constructing the banking infrastructure needed to access agents in market II. This comparative-advantage assumption is rooted in the notion that operating firms specialize in real production, while banks specialize in intermediating between markets I and II, with neither possessing the ability to be a low-cost competitor in areas of non-specialization.

In our model, then, intermediation is (socially valuable) arbitrage between markets I and II. It is socially valuable because a liquidity premium reflects the potential gains from trade by using the safe collateral constructed from assets in market I to meet the demand for safe debt in market II.

Our conclusions about bank capital structure depend critically on the assumption that such arbitrage cannot be conducted at zero cost. Costless arbitrage would fully integrate markets I and II into one perfect/complete capital market. The existence of a liquidity premium means that agents in II pay a higher price for a safe claim than the price that prevails for an identical (riskless) claim in I. This pricing differential would be eliminated by costless arbitrage between markets I and II, and our model rules out such arbitrage so that there is an economically meaningful role for intermediation.

Specifically, we assume there are scale-related costs of banking that strictly determine the size of each bank and that limit the ability of all banks to generate value by producing safe debt claims for agents in market II backed by asset collateral obtained in market I. These costs represent outlays needed to capture the gains from trade by bridging the gap between markets I and II. They also effectively serve as arbitrage impediments that prevent equalization of the rate of return on safe bank debt (deposits) issued to parties in market II with the rate of return on safe portfolios constructed from assets obtained in market I. The reason a rate of return differential persists is that it must be large enough to cover the costs for banks of bridging the gap between the two market segments. We discuss these costs and their impact on bank capital structure in section 5.

Although it is not required for our conclusions about bank leverage and risk management, our model can also accommodate agents that value opportunities to borrow, but do not have access to market I, and instead must transact in market II's imperfect capital market. Banks now serve a second useful purpose by screening credit risks and extending loans. They can generate value from the difference between what they earn on loans extended to borrowers in market II and what they would have to pay for an equivalent risk asset under I's perfect market conditions.

We use this simple segmented-markets model structure to incorporate – through the agents in market II – a demand for liquid claims per se. This approach seems reasonable given that many individuals and businesses do not participate directly in the stock or bond market, but do hold cash balances in bank and money market accounts. Use of a segmented-markets model structure has precedent in the intermediation literature, most notably in Merton's (1990, p. 441) pioneering analysis. More recently, Allen and Carletti (2013) employ such a structure to model bank capital structure. On a more general level, Titman (2002) argues that segmented-markets models – which sustain deviations from MM through arbitrage impediments in the spirit of Shleifer and Vishny (1997) – can explain many otherwise puzzling features of corporate financial policies.

Our segmented-markets structure is not the only approach that is consistent with a liquidity premium (rate-of-return discount) on safe bank debt. For example, one could assume there are liabilities produced by banks – with special features referred to as liquidity – that economic agents value because they provide safe and easy access to resources in a way that other financial instruments cannot. These economic agents are willing to hold these liabilities because of their liquidity advantage, even if they have a lower pecuniary return than other financial instruments. This approach is akin to including money in the utility function, a model structure that is common in monetary economics and that is used by Stein (2012) in his analysis of bank liquidity production. With this approach, a market premium for bank debt is traceable to the liquidity demand that arises separately from the traditional portfolio demand for assets.

With either approach, high leverage is optimal for banks when there are agents who willingly pay a premium for liquid claims (bank debt) that provide immediate, assured access to capital. In section 4, we show that a bank's capital structure implies high leverage assuming (i) bank scale is fixed, (ii) liquidity commands a premium, and (iii) the bank can exploit access to market I when optimizing its asset structure for liquid-claim production. In section 5, we treat bank scale as endogenous and show that, when banks produce socially valuable liquid claims, there are two types of equilibria – one in which liquidity is priced at a premium and another in which it is not.

The focus in section 5 is on the empirically more relevant case in which liquid claims are priced at a premium. [See Krishnamurthy and Vissing-Jorgensen (2012a, 2012b) for evidence that Treasury security prices embed a liquidity premium, i.e., there is a market premium for safe/liquid assets.] We show that such an equilibrium obtains in our model given asset-side costs of banking, e.g., of operational infrastructure, risk management, etc., provided that the demand for liquid assets is high enough. In this equilibrium, bank scale is strictly determinate and individual banks are highly levered.

Allen and Carletti (2013) develop a segmented-markets model in which short-term bank debt is differentiated from other sources of funding. In their model, bank capital structure matters in the presence of bankruptcy costs, but the MM leverage irrelevance result applies when such costs are zero. In our model, there are no bankruptcy costs, and bank capital structure choice matters because it is through safe debt issuance that banks generate greater value when liquidity is priced at a premium. Also, with agency costs ruled out, the attraction of bank leverage in our model is not due to favorable incentive effects of debt, e.g., as in Calomiris and Kahn (1991), Flannery (1994), and Diamond and Rajan (2000).

Consistent with Diamond and Dybvig (1983) and Gorton (2010), we define a perfectly liquid financial claim to be one whose value is not sensitive to the arrival of new information. Such a claim provides assured access to capital in the intuitive sense of a riskless security that provides its owner the same amount in every state of the world. As Diamond and Dybvig emphasize, the demand for such claims reflects uncertainty and the prospect that future liquidity shocks (arrival of new information) will dictate a need for funds for the party seeking liquidity. This general view of liquidity as a valuable asset has a venerable history (see, e.g., Keynes (1936) and Tobin (1958)).

Given a premium for safe/liquid claims, we will show that high bank leverage is optimal in our model. However, the optimality of high leverage is *conditional* on a bank optimizing its asset structure to create suitable safe asset collateral to support abundant safe debt, as we detail next.

3. Risk management for maximal liquid-claim production

In our model, banks can make loans with risky payoffs and they can purchase a wide variety of (risky and safe) financial assets because they have unfettered access to perfect capital markets – market I – just as in the perfect/complete market settings that yield the MM theorem. Consequently, in our model the investment opportunities of banks do not differ in any way from the investment opportunities that are typically assumed in the literature. In particular, the assumption that banks have access to a perfect capital market is implicit in prior studies that use debt-equity neutrality as the baseline to argue that mandated leverage reductions would not raise the social cost of funding banks.

Although this capital-market access assumption is commonplace, it has an important implication that has not been previously recognized: The perfect markets conditions that seemingly make MM applicable to bank *capital* structure actually enable banks to engage in risk-management strategies that create safe *asset* structures. In our model, banks can and will use such strategies to support large amounts of safe debt and capture the premium paid for liquid claims.

To see why banks in our model find it optimal to construct asset portfolios that are riskless, suppose a bank has hedged all of its loans and other assets to obtain such a portfolio. This is always possible when the bank has access to a complete market, or access to an incomplete market in which banks can create a riskless portfolio through suitable diversification and hedging. Now consider a hypothetical portfolio restructuring in which the bank sells off one dollar of its riskless asset holdings and uses that dollar to buy any other risky claim available in market I. Since there are no arbitrage opportunities *within* that perfect market segment, the purchased claim must have a lower payoff than the sold riskless claim in at least one state of nature.[3] If the bank made this risk-increasing portfolio adjustment, it would now have a lower capacity for producing safe debt claims. If, in equilibrium, a bank has the incentive to maximize production of safe claims, then its value would be lower because the now-risky portfolio supports less safe debt. In the next section, we show that a bank maximizes production of safe claims when such claims command a premium.

Intuitively, then, even if the loans that a bank originates are quite risky, a bank with access to perfect markets can and will undertake security transactions to convert its overall risky portfolio into one with no risk. In our model, this portfolio transformation is the bank's optimal *asset*-side policy when safe/liquid claims (bank deposits) are priced at a premium. By engaging in risk management to transform a risky portfolio into a riskless portfolio, banks arrange collateral support for capturing the greatest possible value from the production of liquid financial claims.

4. Optimal capital structure with bank scale fixed

This section follows the standard MM approach of analyzing capital structure with investment policy (bank scale) fixed. Section 5 adds asset-side costs of banking and treats bank scale as endogenous.

We use an infinite-horizon (t = 0, 1, 2, etc.) model as in MM (1961), with modifications to include a potential premium paid for safe/liquid financial claims delivered to market II. We analyze a given price-taking bank that has exploited the MM capital-market conditions to obtain a safe asset mix to foster liquid-claim production. The bank can issue equity or safe/liquid debt claims (deposits), which can be rolled over in perpetuity. With only safe debt in the capital structure, there is no chance of bank failure

[3] Note that the impediments to arbitrage in our model apply *across* markets I and II, not *within* the perfect/complete structure of market I. Safe debt claims supplied by intermediaries to those willing to pay for liquidity are constructed from securities in the perfect-market segment I where safe asset portfolios command a higher return than the bank pays to depositors in market II on the safe debt it creates from its asset collateral. [The same asset-liability return comparison holds when the model include bank borrowers with impaired access to market I since the risk-equivalent returns on bank loans are no lower than in market I.] Such a rate-of-return differential could not survive absent costs that impede arbitrage across markets I and II. In our model, scale-related costs of banking create such an impediment.

and therefore no possibility of runs that could lead to bank losses.

The natural interpretation is that the bank's capital-structure choice is among different mixes of liquid claims – immediately redeemable safe debt – and equity financing. To the extent that long-term debt issued by banks does not provide liquidity services, the MM (1958) analysis would apply to the choice between long-term debt and equity. However, if banks' long-term debt is sufficiently safe to yield liquidity benefits, then the arguments of Gorton and Pennacchi (1990) imply that banks would find safe long-term debt to be superior to equity as a funding tool.

For simplicity, we take the capital market's one-period rate of interest, r, to be constant and assume the same is true for θ and ϕ, which are defined as:

θ = "liquidity spread" or rate-of-return discount that those purchasing liquidity from banks accept in exchange for assured access to capital.

ϕ = "loan spread" or rate-of-return premium paid on bank loans by those with limited access to capital markets.

Our model formulation is compatible with, but does not require, synergies between bank lending and the production of liquid claims, e.g., as discussed by Kashyap, Rajan, and Stein (2002). In particular, our conclusion about the optimality of high leverage in no way depends on the bank earning a positive spread by making loans at a rate-of-return premium, as can be verified by setting $\phi = 0$ in all that follows. In principle, ϕ is a certainty-equivalent parameter that is bank specific, since it depends on the risk structure of the loans that a bank extends. We keep the notation simple and avoid indexing ϕ by bank since our capital structure conclusions hold for all values of ϕ.

A positive liquidity spread ($\theta > 0$) is essential for banks to have a strict incentive to lever up. This incentive is present only to the extent that the bank's asset collateral supports safe debt. Levering up with risky debt does not capture a market premium for liquidity. We provisionally treat θ as parametric to a given bank, but in section 5 discuss how $\theta > 0$ can arise in equilibrium.

The asset side of the bank's balance sheet reflects its purchases of securities at a fair price to earn the rate r (when converted through risk management to a safe asset portfolio) and its extension of loans that yield the return $r(1 + \phi)$ (again when optimally converted to a safe portfolio). These assets collectively serve as collateral whose returns are used to pay interest on short-term debt and dividends that distribute the bank's free cash flow (FCF) to its shareholders. [In section 5, FCF also nets out the costs of providing banking services. Those costs make bank scale determinate and that, in turn, induces banks to distribute FCF as it is earned.] As discussed in section 3, while the individual assets held by the bank can be risky – and surely are for loans – the bank's portfolio of assets is not risky at the optimum. For each dollar of debt the bank issues at a given date, it pays $r(1 - \theta)$ at the next date, i.e., one period forward in time.

The return on a bank's assets is always sufficient to pay the interest on its debt. This is production of liquidity in the purest sense. It provides the purchaser of a debt claim with 100%-assured access to capital at any future date. When $\theta > 0$ and $\phi = 0$, we have MM with one new feature: The existence of a demand for liquidity that results in a market premium on safe debt that is paid by those who seek assured access to capital. That demand is filled by the production of safe debt claims by banks.

Let I denote total bank assets at $t = 0$ when the bank is formed. The same asset level is maintained at each future date $t = 1, 2, 3$, etc. Consistent with MM, the bank's investment policy is fixed, with all FCF distributed to equity as it is earned. To analyze the bank's capital structure choices, i.e., the funding choices that affect the liability side of the bank's balance sheet, we further define:

x	= fraction of capital at $t = 0$ raised by issuing debt ($0 \leq x \leq 1$).
$(1 - x)$	= fraction of capital at $t = 0$ raised by issuing new equity.
$D = xI$	= value of debt (created at $t = 0$ and rolled over in perpetuity).
$(1 - x)I$	= value of equity contributed at $t = 0$.
z	= fraction of capital invested in loans that yield $r(1 + \phi)$.
$(1 - z)$	= fraction of capital invested in securities purchased in the capital market to yield r.

The requirement that $x \leq 1$ (equivalently, $D \leq I$) precludes the bank at all dates from issuing debt above the level of assets and using the excess resources to fund immediate payouts. If we instead allow $x > 1$, the bank could accelerate payout of the present value of the FCF stream. In that case, the highest feasible value of x depends on the PV of the FCF stream (defined below) optimized for maximal debt issuance (and that bound is an increasing function of θ). Banks would then push D above I as far as possible when $\theta > 0$. As discussed in section 5, when we add scale-related costs to the model, banks do not push D to as high a level, but their optimal leverage ratios will nonetheless still be high because those costs will be capitalized into, and thus diminish, the value of equity.

For all banks, we require $0 \leq z \leq 1$. Within these bounds, higher values of z imply greater bank value through the capture of the loan premium ϕ. We treat z as parametric and allow different banks to face different loan ceilings ($z < 1$) due to differences in their credit-evaluation abilities: Banks that are more efficient at credit evaluation extend a larger volume (zI) of loans earning ϕ.

Treating z as parametric makes no difference for this section's basic capital structure analysis, which shows that high leverage is optimal for all z and ϕ when $\theta > 0$. Section 5 shows that banks that are more efficient at credit evaluation (i.e, those with higher z-values) are better able to compete (with unregulated shadow banks) when regulators impose ceilings on their leverage ratios.

At each future date $t > 0$, the bank's free cash flow (FCF) equals its cash inflow from loans plus its cash inflow from capital market securities minus the interest it pays on its debt:

$$FCF = r(1 + \phi)Iz + rI(1 - z) - r(1 - \theta)xI = [1 + \phi z - (1 - \theta) x]rI \tag{1}$$

Note that FCF is the bank's net interest margin in dollar terms. It is the residual cash flow owned by shareholders. It does not include a charge for any equity raised when the bank was initially capitalized at $t = 0$. The FCF term is each period's total dollar return to all equity, including any contributed capital at $t = 0$. [We show below that, when $\theta > 0$, all funding come from debt, so that positive values of equity are "franchise" values from liquidity and loan spreads, net of the scale-related costs of banking.] In operating firms, FCF excludes financial policy flow variables. Banks are different because financial flows are the inputs and outputs they utilize to generate value for their shareholders.

The value of bank equity, E, at $t = 0$ is the discounted value of the FCF (and dividend) stream:

$$E = FCF/r = [1 + \phi z - (1 - \theta)x]I \tag{2}$$

The current (initial) shareholders' wealth at $t = 0$ is $W = E - (1 - x)I$, which nets out the value of any capital contribution they make. Substitution of (2) into the shareholder wealth expression yields:

$$W = [1 + \phi z - (1 - \theta)x]I - (1 - x)I = [\phi z + x\theta]I \tag{3}$$

MM (1958) show that, holding investment policy fixed, capital structure has no effect on value. From (3), in our model, the value impact of changing leverage while holding investment policy fixed is:

$$\partial W/\partial x = \theta I \tag{4}$$

The MM result holds here when $\theta = 0$, since then (4) implies $\partial W/\partial x = 0$ for all x ($0 \leq x \leq 1$).

The MM theorem does not hold when $\theta > 0$. Now, the optimal financing mix is $x = 1$ because $\partial W/\partial x = \theta I > 0$ for all x. Debt dominates equity for any investment scale, I, because of the spread earned by borrowing at a rate that nets out the liquidity premium, θ. There is no spread earned by issuing equity.

The bank's incentive to issue debt depends on θ, and not on ϕ or z. ϕ and z affect the asset side of the bank's balance sheet, and thus affect the scale of collateral used to produce liquid claims. Consequently, $x = 1$ is the unique optimum regardless of the values of ϕ and z. Higher values of ϕ and z raise the value of equity. That has an indirect effect on the bank's leverage ratio at the optimum, as detailed immediately below. However, it does not diminish the bank's incentive to maximize the issuance of liquid claims (set $x = 1$) conditional on its asset structure when $\theta > 0$.

The key implication of our analysis is that the bank's optimal capital structure maximizes *safe* debt issuance against its *safe* asset collateral. This translates to high leverage *only if* the bank has arranged its asset structure to support abundant amounts of safe debt. Absent such asset-side risk management, banks have no incentives to lever up. The reason is that, in our model, banks can capture a liquidity premium only for safe debt. Once debt becomes risky because the underlying asset collateral is risky, there is no premium paid for debt and no incentive to increase leverage further. Hence our argument is not a liquidity-based re-working of the MM corporate tax model, which has a strict incentive to lever up to the maximum extent possible, even when a firm has a highly risky asset structure.

The optimal leverage ratio (based on the values of D and E when $x = 1$ and $\theta > 0$) is:

$$D/(D + E) = 1/[1 + \theta + \phi z] \tag{5}$$

This leverage optimum reflects the fact that equity has a positive value equal to the present value of the FCF stream generated by issuing debt to capture the positive liquidity spread $\theta > 0$ (and by extending credit to capture the loan spread when $\phi z > 0$). Equity value is $E = [\theta + \phi z]I$, and debt value is $D = I$, when the bank sets $x = 1$. This says that the bank generates value for shareholders from the sum of the liquidity and loan spreads that it earns.

With $\theta > 0$, all bank funding comes from debt issuance. A positive value for equity reflects a positive "franchise" or "charter" value, not contributed equity funding. The leverage ratio in (5) takes on values less than one solely because of a positive franchise value to equity. In section 5, we discuss franchise value and optimal leverage in the face of scale-related costs of banking (which are excluded from the current analysis which holds bank scale fixed). In that case, optimal leverage is not given by (5), but the general properties of optimal leverage are unchanged, as discussed in section 5.

As the liquidity premium, θ, declines, optimal leverage increases (see (5)) due to the fall in FCF from liquidity production, which erodes equity value. The leverage increase is due solely to a decline in franchise value. It is not due to incentives to issue more safe/liquid debt. For all positive values of θ, the optimal level of bank debt, D, is at the same corner solution when scale is fixed. [A similar corner solution obtains in the MM corporate-tax model: With any positive tax rate and investment policy fixed, firms add debt to whatever extent they can, and higher tax rates do not affect the corner-solution optimum.] In our model, since bank shareholders capture the premium from liquid-claim production, a ceteris paribus increase in θ implies a higher value of equity, E, and lower leverage, per (5).

Optimal bank leverage is high given empirically plausible values for the model's parameters. To see why, examine (5) and note that one would as an empirical matter expect θ and ϕz to be small positive numbers. Huge liquidity premiums ($\theta >> 0$) or huge loan spreads ($\phi >> 0$) seem implausible as an equilibrium property in today's market where shadow banks produce massive supplies of relatively liquid claims and junk bonds are used aggressively as substitutes for bank loans. Hanson, Kashyap, and Stein (2011) apply the estimates of Krishnamurthy and Vissing-Jorgensen (2012a) to argue that a plausible upper bound on θ is 0.01.

We can think of very low levels of θ as capturing market settings with strong competition among producers of liquidity. The development of financial-engineering tools and, more generally, of shadow banking – including but not limited to money market funds – implies downward pressure on liquidity spreads (θ), which according to (5), implies upward pressure on optimal bank leverage.[4]

The well documented increase in bank leverage since the early 1800s could thus be explained by a long-term trend toward greater competition in the supply of liquid claims. Upward pressure on bank leverage from advances in financial engineering and shadow banking was plausibly reinforced by the development of the junk-bond market and other such innovations. These developments likely put downward pressure on loan spreads, ϕ, which (per (5)) also leads to higher optimal bank leverage.

Gorton (2012, chapter 11) summarizes the U.S. and international evidence of the long-run evolution toward higher bank leverage ratios. He discusses a broad variety of institutional changes that plausibly contributed to this trend, including changes in bankruptcy laws and technological improvements in portfolio management. He also discusses how competition from money-market funds and junk bonds eroded bank profitability in the 1980s (see especially pp. 126-129).

In our model, high bank leverage generates no systemic risk. This result is an implication of our assumption that banks have access to perfect/complete market I, which is another way of saying that MM's capital-market conditions are operative for banks. With access to market I, banks can construct perfectly safe asset portfolios. They will do so because such portfolios create maximal ability to produce (socially and privately) valuable liquid claims. These liquid claims are riskless in our model, and so there is no chance of bank default or systemic meltdowns triggered by bank defaults. The latter property is not a radical departure from the MM benchmark case: Under MM there would not be systemic meltdowns even if banks had risky asset portfolios because default never creates social costs in that basic setting.

Banks always have incentives to maximize liquid-claim production against whatever safe asset collateral they have. That is how they capture the greatest value from liquidity production. Suppose for the moment that a bank faces risk-management costs that impede the attainment of a perfectly safe asset

[4] In section 5's analysis, spreads reflect the scale-related costs faced by the marginal entrant into banking. We can think of the factors discussed in this and the next paragraph as lowering costs for the marginal entrant. For any $\theta > 0$, the marginal entrant must lever to the hilt just to survive in section 5's equilibrium. Infra-marginal banks will have leverage ratios below one that reflect the positive "franchise value" of their equity (which is dictated by their cost advantage over marginal entrants). Increased competition will squeeze those franchise values and lead to higher leverage ratios for infra-marginal banks.

structure. Bank equity is, of course, no longer riskless when there is even the tiniest amount of asset-side risk. A model with a small risk-management impediment is arguably superior to the perfect risk-management formulation we emphasize. Risky equity is more realistic, and it is clearly differentiated from safe debt in terms of liquidity properties. In the end, it makes no difference for our conclusions if we posit small risk-management costs, as the perfect risk-management formulation is simply the limiting case obtained as risk-management impediments become arbitrarily small.

Even with imperfect risk management, $\theta > 0$ and (4) together imply that the bank still has the incentive to issue safe debt to the maximum extent possible, where that maximum is dictated by the left-tail properties of its now-risky asset portfolio. The key distinction of this modified model is that there is now a role for equity as a cushion to support safe debt issuance. A more complete analysis would also recognize that having larger amounts of equity in the capital structure can offer other benefits. For example, Holmström and Tirole (1997) and Mehran and Thakor (2011) argue that more bank equity leads to better monitoring of borrowers. Sections 5 and 6 address the implications of the existence of limitations on the efficacy of risk management.

Many papers argue that banks benefit from high leverage because it maximizes the value of the put option they have against a deposit-insurance fund. In our model, there is no deposit insurance, and so there is no put-option motive for high leverage.

There is also no need for bank leverage limits, as they yield no social benefits in our model. The reason is that banks can fully hedge asset risk in market I, i.e., they have access to perfect/complete capital markets as in models that yield the MM theorem. With such capital-market access and a liquidity premium paid for delivering safe debt claims to market II, banks optimally choose safe asset structures, and so there is no default risk and therefore no systemic risk. Hence, there is no systemic-risk reduction from regulations that place a cap on bank leverage.

In our model, when liquid claims command a premium, shareholders would strictly prefer that the bank not issue equity to pay down debt, and they would not care if the bank issued equity to fund dividends or share repurchases. If a bank were forced to raise equity and use the proceeds to retire safe debt, shareholders would be worse off (without any social gain) due to the decline in the dollar value of the liquidity spread that they capture. If the bank used the equity proceeds to fund equity payouts, current shareholders would obviously be neither better off nor worse off (per MM (1961)). Total equity and total debt would be unchanged, as would be bank leverage, scale, and value.

Finally suppose that the bank is forced to raise equity and that it invests the proceeds in risky assets (with limited liability) that are priced to earn a normal return. There is no change in the dollar amount or the risk of the bank's debt, which is still safe because the new assets have limited liability. Bank leverage is now lower because bank scale is larger. When bank scale matters, shareholders will be worse off if this expansion moves the bank away from optimal scale.[5] Since the bank's debt was already safe, there would be no systemic benefit from forcing the bank to scale up with equity, only a cost. If the expansion moves the bank toward optimal scale, shareholders will of course be better off. However, the new lower leverage ratio is strictly sub-optimal, as a debt-funded expansion would capture an incremental liquidity premium and the benefit of operating closer to optimal scale. The best overall strategy for the bank is to operate at optimal scale while issuing as much safe/liquid debt as can be supported at that scale, as detailed in section 5.

5. Optimal bank scale, capital structure, and liquid-claim production

We next analyze optimal bank capital structure when (i) liquid claims command a premium ($\theta > 0$) and (ii) there are costs of intermediation that make bank scale strictly determinate and that serve as impediments to arbitrage that allow a liquidity premium to obtain.

[5] Admati and Hellwig (2013, p.149) argue that one can reduce bank leverage at zero social cost by holding debt – socially valuable liquid claims – constant and increasing assets and equity. There is no social cost to this strategy when bank scale is irrelevant. But when scale matters, mandated expansions can force banks away from optimal scale, thus damaging their cost efficiency. Alternatively, if each bank operates at optimal scale but is forced to have less safe debt than its assets can support, there will be more banks in existence than necessary – incurring more costs in total than necessary – to produce the same amount of liquid claims. The argument also assumes that bank scale expansions would not trigger material additional TBTF costs (Admati, DeMarzo, Hellwig, and Pfleiderer (2011, pp. 50-51)). This is open to debate. Bank assets would need to expand by 29% in order for the debt-to-assets ratio to decline from 97% to 75%. The banking sector would be 4.5 times its current size if asset expansions were used to reduce bank leverage from 90% of total assets to the 20% ratio that characterizes the typical industrial firm.

There are two possible equilibria in our model: $\theta > 0$ and $\theta = 0$. In both cases, aggregate bank debt is strictly determinate and tailored to service the demand for liquid claims from parties with impaired access to capital (i.e., market II described in section 2). When $\theta = 0$, the equilibrium is analogous to Miller (1977), with leverage irrelevant for any one bank. When $\theta > 0$, every bank has an incentive to lever up to capture the premium for producing liquid claims.

The key condition for $\theta > 0$ is that liquid-claim demand is large relative to the efficient size of the banking sector. In this case, all bank funding comes from safe/liquid debt. Optimal leverage (D/(D+E)) is generally high, but it is less than 100% debt for infra-marginal banks whose equity values are positive due to the franchise value from the cost advantages they have over marginal entrants into banking.

5.1 The aggregate supply of safe/liquid claims is bounded (finite)

There is a natural instinct to think that, when banks have access to perfect capital markets in the MM sense (market I), safe/liquid claims can be produced in unlimited quantity at zero cost, thereby always dictating a zero price for liquidity. Not so.

Even when banks can access perfect capital markets, the aggregate supply of safe debt claims is bounded (finite) and can fall short of the aggregate demand for such claims when there is no market premium for liquidity ($\theta = 0$). To see why, consider a simple example in which each operating firm generates value only in a single state of nature so that the securities that it issues are Arrow-Debreu state-contingent claims. No operating firm is technologically able to supply securities to meet the demand for liquid claims from those with imperfect access to markets. Meeting that demand requires the aggregation of securities over many firms to create riskless claims. That is what banks do through their risk-management and capital structure policies, subject to the bounds on the aggregate production of liquid claims implied by decisions on the real side of the economy.

The ability of banks to arrange their asset portfolios to support the creation of safe debt claims is bounded by the aggregate resources available from the real economy in the state of the world with the lowest resource total. [Here, the aggregate resources in the real economy refer to the total payoff from all securities in market I plus the payoff from bank loans made to parties in market II, when the model is expanded to include bank borrowers in market II.] It is impossible to create riskless claims in quantities beyond the level of resources in the state with the worst payoff. With sufficient demand for liquid claims by parties with impaired access to capital, supply cannot satisfy demand when $\theta = 0$. The result is upward pressure on θ to ensure that liquidity is allocated to those who value it most highly.

The same rationing problem exists when we consider operating firms that issue conventional debt and equity securities, not simple Arrow-Debreu claims. Some of these firms may have sufficiently safe operating policies that they can issue some high-quality (riskless) debt that provides the holder with assured access to capital. The amount of liquidity they can supply is limited by the risk of their operating policies. If the market price of liquidity is sufficiently high, they will tilt their operating policies toward the production of riskless cash flows, which can support some additional safe debt issuance. We assume diminishing returns to scale in real-economy production, so that increasing the output of riskless claims eventually becomes unattractive, and is therefore inherently limited.

The point is that some operating firms will be attracted by a positive price of liquidity to issue some safe debt. They will compete (as shadow banks) through their capital structure decisions to supply liquid claims. There are technological limits to operating firms competing with banks in this fashion since their real production is inherently risky and therefore poorly suited to support safe debt.

Intermediaries are firms that have a comparative advantage in managing overall asset risk to mitigate such difficulties so that they can generate maximum value from production of safe/liquid debt. Risk-management techniques, including hedging strategies, basic diversification principles, the use of derivatives, and financial-engineering methods, are the tools banks use to reduce risk to support production of greater amounts of safe debt. The banks that survive in equilibrium are those that are most efficient at producing asset allocations to support capital structures that supply liquid claims to those parties seeking assured access to capital.

5.2 A Miller-style equilibrium with $\theta = 0$ in the market for safe/liquid claims

Just for the moment, assume there are no costs of banking. In this case, banks collectively can, at zero cost, extract the full potential of the real economy to generate riskless claims and supply that amount of safe/liquid debt to market II. The aggregate supply curve of such debt is horizontal at $\theta = 0$ until the finite upper bound on supply is reached. The supply curve turns vertical at that upper limit and remains so for any and all $\theta > 0$. If market II's aggregate demand for liquid claims when $\theta = 0$ falls short of the upper bound on aggregate supply, then the equilibrium price of liquidity is $\theta = 0$.

The resultant equilibrium is analogous to Miller (1977), except now liquid-claim demand, not tax heterogeneity, is why aggregate debt matters. The market-clearing quantity of safe debt falls in the horizontal section of the aggregate supply curve at the quantity of liquid-claim demand from market II when $\theta = 0$. Any one bank viewed in isolation is indifferent to having low, medium, or high safe debt at $\theta = 0$. If banks collectively responded to $\theta = 0$ by issuing no or low debt, the aggregate supply of liquid claims would fall short of demand. This would put pressure on the banking sector to produce more safe debt in total until market II's liquid-claim demand is satisfied. Some banks will comply because each one views the quantity of liquid claims that it produces to be a matter of indifference when $\theta = 0$.

In this Miller-style equilibrium, the liquid-claim output of banks in the aggregate matters but, since $\theta = 0$, there is no incentive for any one bank to adopt a high leverage capital structure. In this case, there is a role for bank funding from equity, which has no cost disadvantage when $\theta = 0$.

Now, suppose instead that the liquid-claim demand in market II is so strong at $\theta = 0$ that it exceeds the potential of the real economy to generate safe/liquid claims. The result will be (i) upward pressure on θ to encourage liquidity to flow to those who value it most highly, and (ii) downward pressure on r to induce the real economy to provide a larger supply of riskless claims. [Since we assume diminishing returns to scale in real-economy production, the latter expansion in the supply of safe claims is inherently limited.] The main point here is that, even when there are no costs of banking, there will be economic forces pushing for an equilibrium in which $\theta > 0$ when liquid-claim demand is strong.

5.3 Equilibrium with $\theta > 0$ in the market for safe/liquid claims

In our model, whether θ is zero or positive in equilibrium depends on the extent of demand in market II for liquid claims relative to the efficient size of the banking sector, with $\theta > 0$ arising when demand is strong in relative terms. To see how $\theta > 0$ arises, suppose first that $\theta = 0$ because liquid-claim demand falls below what the banking sector willingly supplies at that price. In gauging what the sector will supply there is, of course, the physical limit on safe debt supply (detailed in section 5.2) that puts a ceiling on supply even when $\theta = 0$. Beyond that basic physical limit, we posit the existence of scale-related costs of banking (detailed in section 5.4) that strictly determine the size of individual banks. These costs imply that no bank has an opportunity for gain through unlimited expansion to produce safe debt when $\theta > 0$.

If we then consider progressive expansions in liquid-claim demand from market II, there will come a point where the aggregate quantity of safe debt supplied to market II by banks at $\theta = 0$ falls short of demand from that market. The result will be upward pressure on θ, increasing the quantity supplied and reducing demand until the two are in balance. Since $\theta > 0$ implies that all external funding of banks comes from safe debt, such an equilibrium has strong liquid-claim demand relative to the efficient size of the banking sector.

The analysis is simpler – equilibrium always has $\theta > 0$ – when banks do not make loans to parties with impaired access to capital. Now, the efficient size of the banking sector cannot exceed liquid-claim demand. Equilibrium requires $\theta > 0$ because, with fixed costs of banking (see section 5.4), $\theta = 0$ would imply that banks could not cover their costs. When banks also make loans, loan spreads can by themselves be large enough to cover the costs of banking. Hence $\theta = 0$ can and, in fact, will obtain unless liquidity demand is strong relative to the efficient size of the banking sector, as explained above.

5.4 Optimal bank capital structure and scale when $\theta > 0$

In what follows, we assume liquidity demand is sufficiently strong that $\theta > 0$, and we explore the implications for the capital structure of individual banks when scale is strictly determinate. Specifically, we let $C(I, z)$ denote the explicit costs of operating a bank, which capture *asset*-side risk-management and operational infrastructure costs. In terms of the costs of intermediation, $C(I, z)$, we intuitively have in mind outlays for (i) operational (physical and contractual) architecture of providing banking services and for (ii) risk management, e.g., identifying the right hedges to neutralize the risks borne by the bank from loans it has extended. In models that go beyond our (intentionally simplified) setup, scale-related costs of banking could also include agency costs and other costs caused by market imperfections. In extended models, there would also be ongoing costs associated with trading securities to maintain suitably hedged asset-side portfolio positions.

Current shareholders' wealth is now given by (3) modified to net out C(I, z), which is denominated in present-value terms:

$$W = [\phi z + x\theta]I - C(I, z) \tag{3a}$$

We assume that C(I, z) includes the present value of a periodic fixed cost component and that marginal costs are positive and increasing in bank scale, I. The costs C(I, z) dictate efficient scale for each individual bank and therefore also for banks in the aggregate. Because these costs deter expansion in the aggregate supply of safe/liquid claims when $\theta > 0$, they effectively impede arbitrage (across markets I and II) that otherwise could force $\theta = 0$. In short, they make it possible to have $\theta > 0$ in equilibrium.

Holding scale fixed at any given level, the marginal impact of increasing debt is still given by (4) so that, with $\theta > 0$, optimal capital structure entails maximal production of liquid claims against the bank's safe asset collateral. The first-order condition for optimal scale is $\partial W/\partial I = [\phi z + \theta] - \partial C(I, z)/\partial I = 0$ (with $x = 1$ since that is the upper bound corner solution when $\theta > 0$). Because marginal cost is positive and increasing in I, the optimal bank scale is strictly determinate, as is the optimal amount of debt.

Bankers often argue that regulatory caps on leverage will damage their banks' ability to compete. Our model indicates there is merit in this view. To see why, suppose there is free entry into banking with all entrants having access to the same technology. Let I* denote the bank scale that minimizes long-run average cost, C(I, z)/I. [Note that, given the fixed cost component of C(I, z), minimum average cost does not occur at infinitesimal scale.]

Given that $\theta > 0$, each new entrant sets $x = 1$ to capture the largest possible amount of value from liquid-claim production. A necessary condition for equilibrium is that the sum of the loan and liquidity premiums are driven down to the point where W is zero at the asset and capital structure optimum:

$$W^* = [\phi z + \theta]I^* - C(I^*, z) = 0 \tag{3b}$$

In equilibrium, the sum of the loan and liquidity spreads $(\phi z + \theta)$ just covers average cost, $C(I^*, z)/I^*$, at the optimal bank scale. Any higher spreads would precipitate entry by new banks that see $W > 0$ at I* and $x^* = 1$. Any lower spreads would precipitate exit because $W < 0$. This is just an application of the standard price theory condition that, with competitive entry by identical firms, market prices adjust until firms are just able to cover long-run average cost.

Now, suppose that "conventional" regulated banks face constraints on leverage that mandate $x < 1$, while "shadow" banks face no such constraints. Equilibrium spreads are set in accord with (3b) by free entry by shadow banks. With regulations that place a cap on leverage in (3a), conventional banks have $W = [\phi z + x\theta]I^* - C(I^*, z)$. This expression is negative given that $x < 1$ is now required, and that condition (3b) describes market equilibrium. Conventional banks will therefore exit the market for liquid claims and be replaced by shadow banks that are not subject to regulatory limits on leverage.

The implication is that conventional banks will not be able to compete with shadow banks that have comparable technologies for liquid-claim production. Conventional banks capture $x\theta$ for each unit of scale, whereas shadow banks capture θ per unit of scale. With the higher payoff to liquidity production, shadow banks just cover average cost at the efficient bank scale. With a lower liquidity payoff and the same technology, conventional banks cannot cover costs.

Conventional banks can offset the disadvantage of regulatory limits on leverage if they are better than shadow banks at loan extension (their ϕz is higher due to higher z) or at the infrastructure and financial-engineering elements of delivering banking services (their $C(I, z)$ is lower).

However, even if conventional banks had such technological advantages, the imposition of regulatory caps on their leverage – but not at shadow banks – will induce a substitution of liquidity production into the unregulated shadow-bank sector. Such a substitution is not unique to our analysis; see, e.g., Hanson, Kashyap, and Stein (2011) and Acharya, Mehran, and Thakor (2013). In general, such a substitution is to be expected when bank leverage matters and capital regulations are binding for some, but not all, banks.

5.5 Franchise value and optimal bank leverage when $\theta > 0$

How high is optimal bank leverage in equilibrium? The marginal entrant into banking will be levered to the hilt. With $\theta > 0$ and with (3b) describing equilibrium, optimal leverage for the marginal bank is $D/[D+E] = 1$ because its equity has zero "franchise" value, i.e., the liquidity and loan spreads are just adequate to cover average cost, which means that $E = 0$ when asset and capital structures are jointly optimized. Indeed, the 100%-debt corner solution is not only optimal – it is the *only* viable choice for the marginal entrant into banking. A bank that is just on the margin of entry will not be able to cover costs if it doesn't lever up to the maximum extent possible at the optimal scale.

For banks that are more efficient than the marginal entrant, equity has a positive franchise value equal to the present value of the difference between spreads and costs. For these infra-marginal banks, $E > 0$ and their optimal leverage ratios are below one. Optimal leverage will still be high in absolute terms, unless the bank in question has an enormous cost advantage over marginal entrants that gives it a correspondingly high franchise value ($E >> 0$).

Even for banks with $E > 0$, all funding still comes from safe/liquid debt because $\theta > 0$ implies that issuing such debt generates value. When $\theta > 0$ in our model, a positive value for the equity of a given bank is due entirely to franchise value, which is traceable to that bank's cost advantage over marginal entrants into banking and/or its superior ability to capture loan-related spreads.

The statement that optimal bank leverage is high describes a property of our model of intermediation, which excludes a variety of real-world factors. As discussed in section 6, consideration of these other factors can plausibly imply that lower levels of leverage than in our model make good sense for real-world banks. To be very clear, we are *not* advocating (or predicting) that the marginal entrant into real-world banking should have a 100%-debt capital structure. Rather, our point here is to highlight the basic principle: High bank leverage coupled with (and made possible by) low asset risk is the baseline expected in an idealized world in which banks have a meaningful role in producing liquid financial claims.

6. Bank production of liquid claims that are nearly, but not perfectly, safe

Real-world banks do not produce liquid claims that provide 100%-assured access to capital with no information sensitivity, as they do in our model. Bank deposits that are nearly, but not perfectly, safe can still provide liquidity benefits and command a premium that encourages banks to have higher leverage than most operating firms. This can arise simply because perfectly safe liquid claims are viewed as too expensive, so that bank depositors prefer (suitably priced) liquid claims that pay off fully in almost all states of the world. It can also arise from a variety of behavioral factors. For example, as Gennaioli, Shleifer, and Vishny (2012, 2013) emphasize, reaching-for-yield behavior by agents who neglect or mistakenly assess risk can lead banks to produce risky "near-money" instead of safe debt.

In our model, a liquidity premium is paid only for perfectly safe debt, and so banks focus on hedging strategies that eliminate asset-side risk because doing so builds the maximum capacity to issue debt that is priced to include a liquidity premium. There is a similar incentive to build the capacity to issue nearly safe debt when such debt commands a liquidity premium. Formally, $\theta > 0$ would apply to riskless and low-risk liquid claims, with θ declining as debt becomes more risky and reaching zero when repayment is sufficiently uncertain that it is not valued for liquidity per se. Now, banks will use risk-management strategies to dampen, not eliminate, asset-side risk and to create the capacity to produce abundant quantities of nearly safe financial claims that capture a liquidity premium.

The provision of liquid claims that are less than perfectly safe plausibly also reflects the inherent difficulties of constructing comprehensive hedges that eliminate all sources of risk that could prevent a bank from fully paying off its liquid claims in every state of nature. Such risk-management difficulties do not exist in the perfect/complete market settings that yield the MM theorem. But they surely exist in the real world where the relevant risks for liquid-claim production include "unknown unknowns" that govern left-tail outcomes and that are not easily integrated into conventional finance models.

Under these more realistic demand- and supply-side conditions, banks still have incentives to control, albeit imperfectly, the riskiness of their asset structures to foster liquid-claim production. The better they can control asset-side risks, the greater their ability to lever up and capture a liquidity premium. The main overall implication is: Banks have incentives to lever up to capture a premium on liquid claims to the extent that risk management effectively limits the left-tail risks of their asset portfolios.

What about the "narrow-banking" strategy of buying Treasury securities to serve as collateral to support safe/liquid debt? This strategy is potentially problematic given the evidence of Krishnamurthy and Vissing-Jorgensen (2012a, 2012b) that Treasuries are priced at a liquidity premium. Such pricing attenuates (and possibly eliminates) the return spread from buying Treasuries that a "narrow bank" can use to help cover its costs. This reasoning suggests that bank lending can play a complementary role in liquid-claim production. Banks that are efficient at screening credit risks effectively create collateral well suited to supporting capital structures with a large quantity of (relatively safe) debt priced to capture the market's liquidity premium.

Risks of default and systemic meltdown are absent from our model, but they are an inherent feature of banking when (i) liquidity demand applies to relatively (but not perfectly) safe debt, (ii) reaching-for-yield behavior characterizes liquid-claim demand, and/or (iii) purging of left-tail risk from asset portfolios is prohibitively costly. Such risks inevitably exist when these conditions are operative because they imply that risky debt is the only viable way to satisfy liquid-claim demand. Therefore, default and systemic risks associated with bank debt are not prima facie evidence of moral hazard, other agency problems, or of tax motives to borrow, although the latter factors may exacerbate those risks.

As Stein (2012) emphasizes, social costs associated with bank leverage can arise because the production of risky debt comes with an externality – the risk of systemic meltdown – that is not fully priced in the market. The result is socially excessive production of risky liquid claims as banks compete to service the demand for liquidity. Gennaioli, Shleifer, and Vishny (2012, 2013) also discuss the over-production of risky liquid claims and argue that systemic risk arises from imperfect bank diversification coupled with risk-measurement errors, especially correlated mistakes in gauging tail risk.

Regulatory limits on leverage can thus make sense because real-world banks do not fully internalize the costs of system-wide collapse, and so they over-produce risky liquid claims. Our analysis highlights a potential downside that should also be weighed in a regulatory cost-benefit analysis: Leverage limits could impair liquid-claim creation by relatively efficient producers, while shifting production to shadow banks and, in so doing, preserve or perhaps even exacerbate systemic risk. Hanson, Kashyap, and Stein (2011, pp. 15-16) discuss the danger of shifting liquid-claim production to shadow banks and argue that similar capital standards should apply to similar credit exposures at both banks and shadow banks.

7. Conclusions

High leverage coupled with low asset risk is the baseline optimum expected of banks in a model that has just enough frictions for banks to have a meaningful role in liquid-claim production. This stripped-to-the-basics model excludes taxes, deposit insurance, and all other incentives for banks to lever up or, in fact, to have any debt in their capital structures.

The analysis yields a simple fundamentals-based explanation why banks have higher leverage than most operating firms: Banks generate value by producing safe/liquid debt for parties who willingly pay a premium for assured access to capital, and so they specialize in constructing asset structures that are safer than most operating firms have, which enables them to issue larger amounts of safe debt.

These conclusions are grounded in the premise that safe debt commands a liquidity premium, which gives banks in our model the incentive to construct asset structures that enable them to have capital structures that include abundant amounts of such debt. Empirical support for the existence of a liquidity premium comes from Krishnamurthy and Vissing-Jorgensen (2012a, 2012b) who report evidence that Treasury security prices embed such a premium.

The foundational elements of the model are the existence of (i) an exogenous demand for liquid claims in the spirit of Diamond and Dybvig (1983), (ii) costs of intermediation that are a function of bank scale and that prevent the unfettered arbitrage that would eliminate a liquidity premium from safe debt prices, and (iii) banks' ability to access perfect/complete capital markets (as in models that yield the MM theorem) which they use to hedge asset-side risk and build maximum capacity to issue safe debt.

Since banks in our model choose not to issue risky debt, the model contains no systemic risk and thus no reason for regulatory caps on bank leverage. In models that move beyond our idealized setting to have banks generating systemic risk while over-producing risky near-moneys, there is a case for limiting bank leverage that is not present in our framework. Our analysis is relevant to the debate over bank capital regulation in that it challenges the prominent view among economists that MM's debt-equity neutrality result offers the right baseline for thinking about bank capital structure so that severe limits on bank leverage would be essentially free to society. The latter view is problematic because it dismisses the possibility that such limits might impair the production of socially valuable liquid claims.

Our model of bank capital structure differs, of course, from the MM model in that leverage is not a matter of indifference. However, the implications of our model are fully compatible with the general MM principle that operating policy is the dominant source of firm value. The reason is that risk management that fosters issuance of abundant safe debt *is* the optimal operating policy of intermediaries that specialize in liquid-claim production in our model.

Risk management plays a crucial role in our conclusion that high leverage is the right idealized-world baseline for banks. With access to perfect/complete markets as in models that yield the MM theorem, banks can and will eliminate asset risk to support maximal production of safe/liquid claims. Real-world banks obviously do not have access to MM-style capital markets. Their hedging abilities are clearly imperfect, which makes their asset structures less than perfectly safe. Consequently, their ability and incentive to lever up to capture a liquidity premium is not as great as in our model.

Nevertheless, the qualitative effects we highlight are still relevant for understanding why banks pursue leveraged capital structures, and why risk management is central to their operating policies. Banks have incentives to use whatever (albeit imperfect) hedging technologies they can to construct asset portfolios with maximal capacity to support the relatively safe debt that captures a liquidity premium. Whether or to what extent banks should have high leverage thus depends on the effectiveness of their risk management technologies. When banks' ability to construct asset-side hedges is poor, their capacity and incentive to issue relatively safe debt is correspondingly limited. When more comprehensive hedging is possible and effective, higher bank leverage will be optimal. Either way, there is a fundamentals-based reason why banks have higher leverage than operating firms: As specialists with a comparative advantage in liquid-claim production, they tailor their asset structures to support greater amounts of liquid claims to capture the premium on such claims.

References

Acharya, V., Mehran, H., Thakor, A., 2013. Caught between Scylla and Charybdis? Regulating bank leverage when there is rent-seeking and risk-shifting. New York University working paper.

Admati, A., DeMarzo, P., Hellwig, M., Pfleiderer, P., 2011. Fallacies, irrelevant facts, and myths in the discussion of capital regulation: Why bank equity is *not* expensive. Stanford University working paper.

Admati, A., Hellwig, M., 2013. The Bankers' New Clothes. Princeton University Press. Princeton, NJ.

Allen, F., Carletti, E., 2013. Deposits and bank capital structure. University of Pennsylvania working paper.

Calomiris, C., Kahn, C., 1991. The role of demandable debt in structuring optimal banking relationships. American Economic Review 81, 497-513.

Cochrane, J., 2013. Running on empty: Banks should raise more capital, carry less debt – and never need a bailout again. Wall Street Journal, March 1.

Dang, G., Gorton, G., Holmström, B., 2013. Ignorance, Debt and Financial Crises. Columbia University working paper.

Dang, G., Gorton, G., Holmström, B., Ordonez, G., 2013. Banks as secret keepers. Columbia University working paper.

Diamond, D., 1984. Financial intermediation and delegated monitoring. Review of Financial Studies 51, 393-414.

Diamond, D., Dybvig, P., 1983. Bank runs, deposit insurance, and liquidity. Journal of Political Economy 99, 689-721.

Diamond, D., Rajan, R., 2000. A theory of bank capital. Journal of Finance 55, 2431-2465.

Diamond, D., Rajan, R., 2001. Liquidity risk, liquidity creation, and financial fragility: A theory of banking. Journal of Political Economy 109, 287-327.

Flannery, M., 1994. Debt maturity structure and the deadweight cost of leverage: Optimally financing banking firms. American Economic Review 84, 320-331.

Flannery, M., 2012. Corporate finance and financial institutions. Annual Review of Financial Economics 4, 233-253.

Gennaioli, N., Shleifer, A., Vishny, R., 2012. Neglected risks, financial innovation, and financial fragility. Journal of Financial Economics 104, 452-468.

Gennaioli, N., Shleifer, A., Vishny, R., 2013. A model of shadow banking. Journal of Finance 68, 1331-1363.

Gorton, G., 2010. Slapped By the Invisible Hand: The Panic of 2007. Oxford University Press. Oxford, England.

Gorton, G., 2012. Misunderstanding Financial Crises. Oxford University Press. Oxford, England.

Gorton, G., Pennacchi, G., 1990. Financial intermediaries and liquidity creation. Journal of Finance 45, 49-72.

Hanson, S., Kashyap, A., Stein, J., 2011. A macroprudential approach to financial regulation. Journal of Economic Perspectives 25, 3-28.

Holmström, B., Tirole, J., 1997. Financial intermediation, loanable funds, and the real sector. Quarterly Journal of Economics 112, 663-691.

Holmström, B., Tirole, J., 1998. Private and public supply of liquidity. Journal of Political Economy 196, 1-40.

Holmström, B., Tirole, J., 2011. Inside and Outside Equity. MIT Press. Cambridge MA.

Kashyap, A., Rajan, R., Stein, J., 2002. Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. Journal of Finance 57, 33-73.

Keynes, J., 1936. The General Theory of Employment, Interest, and Money. Macmillan. London.

Krishnamurthy, A., Vissing-Jorgensen, A., 2012a. The aggregate demand for treasury debt. Journal of Political Economy 120, 233-267.

Krishnamurthy, A., Vissing-Jorgensen, A., 2012b. Short-term debt and financial crises: What we can learn from U.S. Treasury supply. Northwestern University working paper.

Mehran, H., Thakor, A., 2011. Bank capital and value in the cross section. Review of Financial Studies 24, 1019-1067.

Merton, R., 1990. Continuous-Time Finance. Blackwell.

Miller, M., 1977. Debt and taxes. Journal of Finance 32, 261-275.

Miller, M., 1995. Do the M&M propositions apply to banks? Journal of Banking and Finance 19, 483-489.

Miller, M., Modigliani, F., 1961. Dividend policy, growth, and the valuation of shares. Journal of Business 34, 411-433.

Modigliani, F., Miller, M., 1958. The cost of capital, corporate finance, and the theory of investment. American Economic Review 48, 261-297.

Myerson, R. Review of Admati and Hellwig's Bankers' New Clothes. University of Chicago working paper, http://home.uchicago.edu/~rmyerson/research/newcloth.pdf

Pfleiderer, P., 2010. On the relevancy of Modigliani and Miller to Banking: A parable and some observations. Stanford University working paper.

Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35-55.

Stein, J., 2012. Monetary policy as financial-stability regulation. Quarterly Journal of Economics 127, 57-95.

Titman, S., 2002. The Modigliani and Miller theorem and the integration of financial markets. Financial Management 31, 101-115.

Tobin, J., 1958. Liquidity preference as behavior toward risk. Review of Economic Studies 25, 65-86.

Wolf, M., 2013. Why bankers are intellectually naked. Financial Times, March 17.