Why High Leverage is Optimal for Banks

Harry DeAngelo and René M. Stulz*

April 2013
Revised August 1, 2013

Abstract

Liquidity production is a central role of banks. High leverage is optimal for banks in a capital structure model in which there is a market premium for (socially valuable) liquid financial claims and no deviations from Modigliani and Miller (1958) due to agency problems, deposit insurance, taxes, or other distortions. This model can explain (i) why bank leverage increased over the last 150 years or so, (ii) why high bank leverage per se does not necessarily cause systemic risk, and (iii) why leverage limits for regulated banks impede their ability to compete with unregulated shadow banks. Although MM's debt-equity neutrality principle does not hold in our model, the model's implications are fully consistent with the MM principle that operating policy is the dominant source of firm value: Creation of a capital structure with abundant (safe) debt *is* the optimal operating policy of financial intermediaries that specialize in the production of liquid financial claims in our model.

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.

Why High Leverage is Optimal for Banks

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 indicates 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 maintain high leverage ratios.

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]

This model treats banks as firms that make loans and assumes that one can ignore any impact on the (private and social) value banks create in their role as producers of liquid financial claims. The idea that liquidity production is intrinsic to financial intermediation is discussed extensively by, among others, Diamond and Dybvig (1983), Diamond and Rajan (2001), Gorton (2010), Gorton and Pennacchi (1990),

---

[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.

and Holmstrom 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 deviates from MM by (i) inclusion of an exogenous demand for liquid financial claims in the spirit of Diamond and Dybvig (1983) and the other pioneering studies referenced immediately above, and by (ii) the existence of costs of financial intermediation that are a function of bank scale. The model allows, but does not require, the existence of some parties that willingly pay a premium to obtain bank loans because their capital-market access is costly or otherwise impaired.

To establish that high bank leverage is the natural (distortion-free) result of intermediation focused on liquid-claim production, the model rules out agency problems, deposit insurance, taxes, and all other distortionary factors. By positing these idealized conditions, the model obviously ignores some important determinants of bank capital structure in the real world. However, in contrast to the MM framework – and generalizations that include only leverage-related distortions – it allows a meaningful role for banks as producers of liquidity and shows clearly that, if one extends the MM model to take that role into account, it is optimal for banks to have high leverage.

Specifically, in our model:

- High leverage is an essential, uniquely optimal feature of bank capital structures when liquidity is priced at a premium due to demand for assured access to capital.
- Banks choose high leverage despite the absence of agency costs, deposit insurance, tax motives to borrow, reaching for yield, ROE-based compensation, or any other distortion.
- Greater competition that squeezes bank liquidity and loan spreads diminishes equity value and thereby raises optimal bank leverage ratios.
- If conventional banks face regulatory limits on leverage while shadow banks do not, the former will be at a competitive disadvantage to the latter. Liquidity production will migrate from regulated banks into the unregulated shadow-banking sector.

- When liquidity is priced at a premium, banks can and will choose safe *asset* structures to support *capital* structures that maximize the production of safe financial claims to satisfy the demand for liquid claims.

- Because the MM capital-market conditions enable banks to construct perfectly safe asset and liability structures, there is no chance of bank default and no chance of systemic meltdown when these conditions hold.

These implications for bank capital structure reflect the fact that a market price premium to induce production of (socially valuable) liquid financial claims is incompatible with debt-equity neutrality. A liquidity premium is a price signal sent to banks and other liquidity producers to supply liquid financial claims that meet the demand for such claims from financially constrained firms and other parties that desire assured access to capital. The liquid financial claims (cash balances and similar claims) held as *assets* by financially constrained parties are drawn from the supply of safe claims produced by the *liability* structure decisions of banks and other liquidity suppliers.

When market prices embed a liquidity premium, bank capital structure decisions matter and are jointly optimized with asset-allocation decisions to capture the maximum value from liquid-claim production. Debt and equity are not equally attractive sources of bank capital. As Gorton and Pennacchi (1990) emphasize, debt has a strict advantage because it has the informational insensitivity property – immediacy, safety, and ease of valuation – desired by those seeking liquidity. High bank leverage is accordingly optimal when the MM model is modified to include a price premium to induce (socially valuable) liquidity production.

Because MM's debt-equity neutrality principle assigns zero weight to the social value of liquidity, it is an inappropriately equity-biased baseline for assessing whether the high leverage ratios of real-world banks are excessive or socially destructive.

Nor does the MM theorem imply that regulatory limits on bank leverage are desirable because they would reduce systemic risk without increasing banks' funding costs. In our idealized setting, such regulations would raise banks' funding costs without changing systemic risk (which is nil because banks' asset structures are optimized under conditions of uncertainty to produce safe financial claims).

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 modified MM model, there are no agency costs or taxes, 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 idealized setting, banks arrange their holdings of loans and securities to construct a safe asset portfolio. Doing so supports greater production of liquid financial claims, which manifests in high-debt capital structures. In contrast, operating firms create value through real project choices, which commonly entail significant cash flow uncertainty.

Capital structure is a sideshow for value creation at operating firms, but it is *the* star (or at least a star) of the show at banks. The risky asset structures of most operating firms are poorly suited to support high leverage or large-scale production of liquidity. In our model, banks exist because specialization and the associated cost efficiencies give them a comparative advantage over operating firms in arranging their asset structures to support capital structures that produce liquid financial claims to meet demand.

The general point is that, given a material market demand for liquidity, 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 to produce liquid financial claims. This is the fundamental reason why debt-equity neutrality is an inappropriate equity-biased baseline for concluding that bank leverage should be curtailed significantly.

The paper thus shows that, if we take an idealized-world model of bank capital structure and include a demand for liquid claims, the right baseline for banks is high leverage, not indifference to leverage. An important caveat is that real-world banks fall far short of the perfect asset-side diversification they adopt under the idealized conditions we study. Therefore, in capital structure models that relax our idealized conditions, 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). The latter papers identify potential dangers of over-production of liquid claims by banks and shadow banks when they create risky near-moneys rather than riskless debt. We discuss this issue in section 5 after we present our main results.

Section 2 describes the MM framework augmented to include socially productive roles for banks in the supply of liquid claims and in the screening and granting of loans. Section 3 derives the optimal bank capital structure in the presence of a market premium for liquidity. Section 4 discusses the equilibrium pricing of liquid claims. Section 5 considers bank leverage and systemic risk when banks produce near-moneys because the production of safe claims is prohibitively costly. Section 6 concludes.

2. MM model with a productive role for banks

We start with the most basic perfect-markets setting in which the MM theorem applies to operating firms and banks. There are no taxes, agency costs, bankruptcy costs, or any other frictions. Nor is there deposit insurance. The capital market is complete, with costless access for all households, operating firms, and banks.[2] Intermediation is redundant because all operating firms and households can directly access capital markets at zero cost. Banks can still exist as "neutral mutations" that generate no benefits, but do no harm. The choice of debt-equity mix does not matter for banks. In this most basic of models, there is no private or social cost to limiting bank leverage.

We modify the basic MM model to allow a productive role for bank lending using the segmented

[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. See footnote 3.

markets approach of Merton (1990, p. 441) in which some agents face significant transactions costs of accessing capital markets, while financial intermediaries do not. Intermediation is no longer redundant: Banks are privately and socially valuable because they screen credit risks and extend loans to parties with constrained access to capital markets. Banks capture value from the spread between what they earn on loans and their cost of capital. All debt-equity mixes are equally costly means of raising funds to make loans priced to earn a premium over the cost of capital. The MM theorem continues to apply to banks, but not to the financially constrained parties who borrow from banks.

We also move beyond the basic MM model by incorporating a demand for liquid assets per se. We do not spell out the micro-level details underlying this demand, and instead note that our conclusions obtain using either one of two approaches, both of which have been used in prior models. The first adopts the segmented-market approach of Merton (1990) and posits the existence of a set of agents that have impaired access to the market, but would like assured access to capital to hedge against liquidity shocks. Such an approach is used, for instance, by Allen and Carletti (2013) in a model of bank capital structure. 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.

The second approach to modeling liquidity demand simply assumes 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 liquid assets 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 because there are agents who willingly pay a premium for liquid claims (bank debt) that provide immediate, assured access to capital. In section 3's partial equilibrium analysis, we consider a single bank facing an *assumed* liquidity premium. This analysis shows that, conditional on any given asset scale, high leverage is optimal because issuing debt is how banks profit from producing (socially valuable) liquidity. The MM theorem no longer applies.

Section 4 shows that a liquidity premium obtains in equilibrium when bank scale is determinate due to asset-side costs, e.g., of financial-engineering, operational infrastructure, etc. Even when bank size is indeterminate due to constant returns to scale, aggregate bank debt is strictly determinate. A liquidity premium obtains with constant returns when liquidity demand exceeds the (finite) upper bound on the supply of riskless debt dictated by the real economy. When demand falls below that bound, the equilibrium is as in Miller (1977), except now taxes are not the reason aggregate debt matters: Any one bank's leverage is a matter of indifference because liquidity is not priced at a premium, but the aggregate supply of liquid claims is strictly determined to meet the demand for (socially valuable) liquidity.

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 short-term 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)). It draws empirical support from Krishnamurthy and Vissing-Jorgensen (2012a, 2012b) who provide evidence that Treasury security prices embed an economically significant liquidity premium, i.e., there is a market premium for safe liquid assets. Given such a premium, high bank leverage is optimal in our model.

Banks in our model optimize their asset holdings to support capital structures that efficiently produce liquid financial claims for parties willing to pay a premium for assured access to capital. In Gennaioli, Shleifer, and Vishny (2013), banks (and shadow banks) also optimize their asset portfolios for the same purpose, but they ultimately face prohibitive costs of forming perfectly diversified portfolios, thus leaving the possibility of systemic meltdown through bad outcomes from correlated tail risk.

In our model, banks make loans with possibly risky payoffs and they can purchase a wide variety of risky securities (as well as riskless claims) because we assume they have unfettered access to capital markets as in perfect/complete-market models that yield the MM theorem. This assumption is not unique to our analysis. It 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 (e.g., see the introduction).

Although this capital-market access assumption is commonplace, it has an important implication that has not been previously recognized: With access to perfect/complete capital markets in the MM sense, a bank can choose a portfolio of assets that is completely riskless.[3] Even if the loans it originates are quite risky, a bank with access to perfect markets can undertake security transactions to convert its overall portfolio into one with no risk. In our model, this portfolio transformation is the bank's strictly optimal *asset*-side policy. The reason is that, by allocating assets to form a riskless portfolio, banks arrange collateral support for capturing the greatest possible value from the production of liquid financial claims.

We are *not* arguing that real-world banks are riskless. Our objective is to characterize optimal bank capital structure when the MM assumptions are modified to allow a meaningful role for liquid-claim production. Section 5 discusses liquidity production for more realistic settings beyond our model's

---

[3] To see why a riskless asset structure is optimal in our model, suppose the bank has traded all of its loan holdings for a riskless portfolio. This is always possible when the bank has access to a complete market, or access to an incomplete market in which a riskless asset (or portfolio of assets) exists among the traded claims. Now consider a hypothetical asset restructuring in which the bank sells off one dollar of its riskless portfolio holdings and uses that dollar to buy any other risky claim available in the market. Since MM's assumptions rule out arbitrage opportunities in the pricing of all claims, the purchased claim must have a lower payoff than the sold riskless claim in at least one state of nature. Consequently, the bank would now have a lower capacity for producing liquid claims. The bank's value would be lower due to its reduced ability to capture the premium for supplying liquidity.

idealized conditions so that banks face prohibitively high costs of eliminating asset-side risk.

3. Optimal bank capital structure

This section follows the standard MM approach of analyzing capital structure with investment policy fixed. We incorporate asset-side costs of banking in section 4, which treats bank scale as endogenous and analyzes the conditions leading to a liquidity premium in equilibrium.

We use an infinite-horizon (t = 0, 1, 2, etc.) model as in MM (1961), with modifications to include liquid-claim production. We analyze a given price-taking bank that has exploited the MM capital-market conditions to obtain a riskless asset mix to foster liquid-claim production. The bank can issue equity or short-term debt, which can be rolled over in perpetuity. The natural interpretation is that the bank's capital-structure choice is among different mixes of liquid claims (immediately redeemable riskless claims) 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 long-term debt has liquidity benefits, as argued by Gorton and Pennacchi (1990), the MM analysis would not apply to banks' long-term debt either.

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 future 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. We provisionally treat θ as parametric to a given bank, but we return in section 4 to a discussion of the conditions on the supply of liquidity that sustain $\theta > 0$ in market equilibrium.

The asset side of the bank's balance sheet reflects its purchases of securities at a fair price in the capital market (to earn the rate r) and its extension of loans that yield the return $r(1 + \phi)$. 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. As discussed earlier, 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 that 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 or money in the purest sense. It provides the purchaser of a debt claim with 100%-assured access to capital in the future. 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 paid by those who seek assured access to capital. That demand is filled by the production of riskless debt claims by banks.

Let I denote total bank assets at $t = 0$. 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. Since we take the asset side of a bank's balance sheet as given, the only choices left for us to analyze are the bank's choices that are related to capital structure, i.e., 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.

This formulation assumes that shareholders only contribute capital and do not receive dividends at $t = 0$ when the bank is formed. 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 $\theta > 0$.

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 4 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.

In each future period $t > 0$, 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:

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

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 equity capital raised when the bank was initially capitalized at $t = 0$. The FCF term is each period's total dollar return to all equity, including the newly contributed capital at $t = 0$. 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:

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

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 liquid claim issuance against its asset collateral. 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.

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 not due to incentives to issue more short-term debt (because $x = 1$ was optimal and remains so as long as $\theta > 0$).

Optimal bank leverage is generally high. 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.

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 evidence for the U.S. and internationally 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 because MM's capital-market conditions are operative for banks. We maintain these conditions to clarify the problems with the widely held view that the MM theorem implies there is no socially productive reason for banks to have high leverage. This view is problematic because the MM capital-market conditions enable banks to construct perfectly safe asset portfolios. Such asset portfolios foster the production of greater quantities of (privately and socially) valuable liquid claims. These liquid claims are riskless in this idealized setting, and so there is no chance of bank default or systemic meltdowns triggered by bank defaults. The latter model property is not a radical departure from the MM benchmark case: Under MM there would not be systemic meltdowns even if banks have 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 transactions and other risk-management costs that impede the attainment of a perfectly safe asset structure so that $x = 1$ is infeasible. With even the tiniest such impediment, bank equity is, of course, no longer riskless. However, $\theta > 0$ and (4) together imply that the bank still has the incentive to produce liquid claims (issue riskless debt) to the maximum extent possible, where that maximum is dictated by the left-tail properties of its now-risky asset portfolio. [The benefit of increasing x shrinks on the margin as θ falls, but it is always positive.] In sections 4 and 5, we return to a discussion of the implications of risk-management costs for bank asset and capital structures.

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.

In this idealized world, there is also no need for bank leverage limits, as they yield no social benefits. Banks optimally choose riskless asset structures and so there is no default risk and therefore no systemic risk. Hence, there is no systemic-risk reduction from regulations that restrict bank leverage.

What if a bank bolsters its equity by selling new shares at a fair price? Bank scale is unchanged if the equity proceeds are used to reduce debt or fund equity payouts. If the bank reduces debt, shareholders are worse off by the decline in the dollar value of the liquidity spread they capture. If the bank uses the equity proceeds to pay dividends or repurchase stock, current shareholders obviously are no better or worse off (per MM (1961)) and both leverage and scale are unchanged.

What if the equity proceeds are invested in securities that earn a normal return? Bank scale increases and leverage decreases with equity issuances that leave the dollar amount of debt unchanged. There is no impact on systemic risk, which is nil both before and after the issuance. However, bank leverage is now sub-optimal (per (4)) relative to its greater scale. Shareholders would be better off if the bank had raised debt instead of equity to capture the liquidity spread, while using the proceeds to create more riskless collateral to support the additional liquid-claim production. [The new bank scale is also sub-optimal when the model is enriched (per section 4) to determine scale, and the bank is initially at its optimum.]

Admati and Hellwig (2013, p.149) note that mandated increases in equity that are used to increase bank assets could leave debt (liquid-claim output) fixed, while reducing leverage ratios and associated distortions. Because this regulatory plan holds bank debt constant at a level that fully services liquidity demand at its pre-regulation level, they conclude that limits on leverage have zero social costs even when bank debt is socially valuable because of its liquidity properties.

This argument is incomplete, however, because it does not consider potential supply-side social costs. For example, Admati, DeMarzo, Hellwig, and Pfleiderer (2011, pp. 50-51) discuss one such social cost: A mandated expansion of equity and assets could induce greater "too-big-to-fail" problems at now-larger banks. They suggest that those TBTF costs could be offset by both (i) dividing banks into smaller entities and (ii) requiring more equity and assets for banks in the aggregate, while holding aggregate debt fixed. Admati et al. also mention that the general issue of efficient bank size is a controversial unresolved topic, but do not pursue the implications of this issue for bank-capital regulation.

The latter issue points toward other (non-TBTF) supply-side costs of the mandated bank equity and asset expansion plan emphasized by Admati and Hellwig. Unless scale is irrelevant, the plan can force banks away from optimal scale, thus damaging the cost efficiency with which they produce liquid claims. 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 quantity of debt required to service liquid-claim demand. Either way, the potential costs of the regulation on the supply-side have to be taken seriously.[4]

The implication is that there are tradeoffs involving impaired liquid-claim production that should be weighed in gauging whether or to what extent regulations mandating larger equity cushions are desirable. Those tradeoffs include many real-world factors (e.g., moral hazard) that our idealized-world model excludes. How the tradeoffs net out for real-world banks is an empirical issue we do not address. Our analysis simply points to interactions among liquid-claim production, leverage, and efficient scale that ought to be considered when regulators weigh the social costs and benefits of limiting bank leverage.

4. Equilibrium in the market for liquid financial claims

There is a natural instinct to think that, when banks have access to perfect capital markets in the MM sense, 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 liquid financial 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, subject to the bounds on the aggregate production of liquid claims implied by decisions on the real side of the economy.

Liquidity is always a scarce asset. The ability of banks to arrange their asset portfolios to support the creation of riskless claims is bounded by the aggregate resources available from the real economy in the

[4] It is instructive to consider this issue in conjunction with Admati and Hellwig's (2013, pp. 7-8) contention that there is no good reason why banks and operating firms have different leverage ratios. The typical industrial firm on Compustat has a debt-to-assets ratio of about 20% whereas, as AH note, many banks have debt-to-assets ratios above 90%. Suppose regulators implement the AH plan to force banks to scale up equity and assets while holding debt fixed. If banks were to match the leverage ratio of the typical industrial firm, the banking sector would be 4.5 times the size of what we currently have. It seems implausible that such a massive increase in the scale of the banking sector could be accomplished at zero social cost.

state of the world with the lowest resource total. 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 debt and equity securities instead of 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. Those cash flows, in turn, serve as asset-side collateral that is capable of supporting a capital structure that supplies safe debt claims to meet liquidity demand.

The point is that some operating firms will be attracted by a positive price of liquidity to enter the intermediation business. They will compete (as shadow banks) through their capital structure decisions to supply liquid claims. There are technological limits to operating firms serving as intermediaries since their real production is inherently risky and therefore poorly suited to support liquidity production.

Intermediaries are firms whose asset allocations are *designed* to mitigate such difficulties. Risk-management techniques, including basic diversification principles, the use of derivatives, and financial-engineering methods, are the tools banks use to create asset structures well suited to support production of liquid financial claims. The banks that survive in equilibrium will be 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.

Suppose that banks collectively are large enough to have extracted the full potential of the real economy to generate riskless claims. Suppose also that they have attained this scale at zero real-resource cost. If the greatest potential aggregate supply falls short of liquidity demand when $\theta = 0$, there will be (i) upward pressure on θ (so that liquidity flows to those who value it most highly), and (ii) downward pressure on r (so that the real economy produces a larger potential supply of riskless claims).

If aggregate demand for liquid claims when $\theta = 0$ falls short of the potential aggregate supply, then the market-clearing price of liquidity is $\theta = 0$. The resultant equilibrium is analogous to Miller (1977), except now liquidity demand, not tax heterogeneity, is why aggregate debt matters. The aggregate supply of liquidity by the banking sector is determinate, and matched to satisfy the aggregate demand for assured access to capital from those parties who value liquidity per se. Any one bank viewed in isolation is indifferent to low, medium, or high debt at the market-clearing price of liquidity ($\theta = 0$).

However, markets will not clear if all banks try to choose low debt capital structures, as the aggregate supply of liquid claims would not satisfy demand. This will put pressure on the banking sector to produce more liquidity. Banks will comply because each one views the quantity of liquid claims that it produces to be a matter of indifference when $\theta = 0$.

A richer model incorporates explicit costs of operating a bank, $C(I, z)$, which capture *asset*-side risk-management and operational infrastructure costs. It is clearly trivial to "show" that bank capital structure matters when there are direct operational costs of producing liquid financial claims ($C(I, z, x) > 0$, with $\partial C(I, z, x)/\partial x \neq 0$). Although such costs plausibly exist, we are *not* invoking them here. We are simply recognizing that there are asset-side costs to operating a bank, which dictate efficient bank scale.

With this richer specification, current shareholders' wealth is now given by (3) modified to net out $C(I, z)$, which is denominated in present-value terms:

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

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 ($x = 1$). The first-order condition for optimal scale is $\partial W/\partial I = [\phi z + x\theta] - \partial C(I, z)/\partial I = 0$. Assuming that $C(I, z)$ is an increasing convex function of I, the optimal scale of each bank is now strictly determinate, as is the amount of bank debt.

At each $\theta > 0$, the supply of liquid claims by any one bank – and by all banks in the aggregate – will be smaller as the marginal cost curve ($\partial C(I, z)/\partial I$) shifts upward. [Optimal scale – and, with it, liquid-claim output – shrinks as the marginal cost curve shifts upward.] When marginal costs are substantial and the demand for liquid financial claims is nontrivial, $\theta > 0$ is required for equilibrium.

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. With $\theta > 0$, each new entrant sets x = 1. Entry by new banks will continue until $\phi z + \theta$ is driven down to the point where W is zero when capital structure is optimized (x = 1):

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

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 $\phi z + \theta$ precipitates entry by new banks that see W > 0 at I* and x = 1. Any lower $\phi z + \theta$ precipitates exit because W < 0 at I* and x = 1.

Now, suppose that "conventional" regulated banks face constraints on leverage that mandate x < 1, while "shadow" banks face no such constraints. Equilibrium θ is set in accord with (3b) by free entry by shadow banks. With regulations that cap leverage in (3a), conventional banks have $W = [\phi z + x\theta]I^* - C(I^*, z)$. This expression is negative given that x < 1 and that the free-entry 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 with x < 1, whereas shadow banks, which set x = 1, 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) or at the financial-engineering and infrastructure 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.

5. Near-money production, bank leverage, and systemic risk

Real-world banks fall short of producing the safe debt that fits our assumed ideal of a liquid claim as one that provides assured access to capital with no information sensitivity. This fact plausibly reflects two considerations excluded from our model, one on the demand side and the other on the supply side.

On the demand side, parties seeking liquidity may be satisfied with (suitably priced) financial claims that are nearly, but not perfectly, safe. Moreover, 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-moneys instead of safe debt. When liquidity demand carries over to relatively safe debt claims, any given bank will have incentives to issue near-money (risky debt) to remain competitive with other producers of liquid financial claims.

On the supply side, the costs of purging risk from bank asset portfolios may reach prohibitive levels. The "narrow-banking" strategy of buying Treasury securities to serve as collateral to support riskless bank debt is problematic given the evidence of Krishnamurthy and Vissing-Jorgensen (2012a, 2012b) that Treasuries are priced at a liquidity premium. Such pricing attenuates (and possibly fully eliminates) the return spread from buying Treasuries that a "narrow bank" can use to help cover its costs. Bank lending, however, can play an important 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.

These more realistic demand- and supply-side conditions do not change the implication that banks find high leverage to be beneficial. The reason is that banks capture a liquidity premium by producing risky near-moneys, just as when they produce riskless debt claims.

Risks of default and systemic meltdown are an inherent feature of banking when (i) liquidity demand applies to relatively (but not perfectly) safe debt, (ii) reaching-for-yield behavior is a material feature of liquid-claim demand and/or (iii) purging of risk from asset structures is prohibitively costly. Such risks inevitably exist when one or more of these conditions is operative because that means that the creation of risky near-money is the only viable way to satisfy liquid-claim demand. Therefore, default and systemic risks are not prima facie evidence of moral hazard, other agency problems, or the effect 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 near-money comes with an externality – the risk of systemic meltdown – that is not fully priced in the market. The result is socially excessive production of near-moneys as banks compete to service the demand for liquid financial claims. Gennaioli, Shleifer, and Vishny (2012, 2013) also discuss the over-production of near-moneys 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 near-moneys. 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 production to shadow banks and argue that similar capital standards should apply to similar credit exposures at both banks and shadow banks.

6. Conclusions

A widely held view among economists is that there is no efficiency-based reason why banks have high-leverage capital structures with debt ratios far above those of non-financial firms.

The model in this paper offers a resolution of this puzzle grounded in the fundamentals of the banking business: Banks specialize in the production of liquid financial claims. They generate value by capturing a liquidity premium from supplying liquid claims (safe debt) to parties who willingly pay the premium for assured access to capital. Banks in our model accordingly optimize their asset holdings (under conditions of uncertainty) to provide safe collateral that supports high-leverage capital structures with abundant quantities of safe debt. In contrast, the asset structures of non-financial firms reflect the generally risky nature of their real project choices, and thus are poorly suited to support high leverage.

Because banks in our model produce riskless debt, the model contains no systemic risk and thus no reason for regulatory limits on bank leverage. In models that move beyond our idealized setting to have banks generating systemic risk while producing risky near-moneys, there is a case for reining in bank leverage that is not present in our framework. Our analysis is relevant to the debate over bank capital regulation only 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 and that, consequently, severe limits on bank leverage would be essentially free to society. The latter view is problematic because it assigns no weight to the possibility that such limits on bank leverage might impair the production of socially valuable liquid financial 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 creation of a capital structure with abundant safe debt *is* the optimal operating policy of financial intermediaries that specialize in liquid-claim production in our model.

Finally, our model differs from standard capital structure models in its assumption that there are material costs of liquid-claim production. This assumption seems reasonable given the findings of Krishnamurthy and Vissing-Jorgensen (2012a, 2012b) that Treasury security prices embed a liquidity premium. If competing banks could produce liquid claims in unbounded quantities at zero marginal cost, the ability to expand supply would purge any liquidity premium from market prices. Conversely, the existence of a liquidity premium in Treasuries supports the view that market prices encourage banks to adopt high-leverage capital structures that supply abundant quantities of liquid financial claims.

References

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.

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., 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.

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.

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

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.