

Staff Working Paper/Document de travail du personnel 2017-7

Banking Regulation and Market Making



by David A. Cimon and Corey Garriott

Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

Bank of Canada Staff Working Paper 2017-7

February 2017

Last updated: February 2020

Banking Regulation and Market Making

by

David A. Cimon¹ and Corey Garriott²

¹Wilfrid Laurier University
Lazaridis School of Business and Economics
dcimon@whl.ca

²Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
cgarriott@bankofcanada.ca

Acknowledgements

For helpful comments and discussions, we are grateful to Gregory Bauer, Elif Dalkir, Amy Edwards, Jean-Sébastien Fontaine, Itay Goldstein, Tamara Gomes, Yesol Huh, Sebastian Infante, Darcey McVanel, Albert Menkveld, Christine Parlour, Joshua Slive, Francesco Trebbi, Kumar Venkataraman, Jan Wrampelmeyer and Liyan Yang. We thank conference participants at the 2017 Paris Finance Meeting, the 2017 FMA Annual Meeting, the 13th Annual Central Bank Conference on the Microstructure of Financial Markets, the 2017 SAFE Microstructure Conference, the 2017 NFA Annual Conference, the 2017 CEA Annual Meeting and the 2017 Bank of Canada–Banco de Espana workshop, as well as seminar participants at the Bank of Canada and Wilfrid Laurier University. We are grateful for the research assistance of James Pinnington. This paper represents the opinions of the authors and does not reflect the views of the Bank of Canada.

Abstract

We model how securities dealers respond to regulations on leverage, position, and liquidity such as those imposed by the Basel III framework. The dealers respond by endogenously moving to make markets on an agency basis, matching buyers to sellers rather than taking client positions on the balance sheet. Agency-based market making creates a cost-risk tradeoff in which investor welfare declines but dealers become less risky. The costs to investors do not show up in all liquidity metrics: while asset prices exhibit greater price impact, bid-ask spreads do not change and trading volumes can even increase, which can help explain the varying findings from the empirical literature.

Bank topics: Financial markets; Market structure and pricing; Financial system regulation and policies

JEL codes: G14; G20; L10

Résumé

Dans notre étude, nous modélisons la réaction des maisons de courtage aux exigences réglementaires relatives aux fonds propres, aux positions et aux liquidités, comme celles du dispositif de Bâle III. Les courtiers réagissent par des mouvements endogènes afin d'assurer la tenue de marché pour le compte de tiers : ils jumellent des ordres d'achat et de vente au lieu de détenir les positions de leurs clients dans leur bilan. La tenue de marché pour le compte de tiers crée un arbitrage entre les coûts et les risques qui réduit le bien-être des investisseurs mais aussi les risques auxquels les courtiers s'exposent. Les coûts assumés par les investisseurs ne sont pas captés dans toutes les mesures de la liquidité. Si les prix des actifs sont davantage touchés, les écarts entre les cours acheteur-vendeur ne changent pas et les volumes des opérations peuvent même augmenter, ce qui pourrait expliquer les conclusions divergentes des études empiriques.

Sujets : Marchés financiers; Structure de marché et fixation des prix; Réglementation et politiques relatives au système financier

Codes JEL : G14; G20; L10

I. Introduction

After the 2007–2008 financial crisis, banking regulators in many countries coordinated on a set of reforms for banking. The reforms, called Basel III, are the third set of banking reforms developed at the Basel Committee on Banking Supervision. Basel III is designed to strengthen the finances of banks by asking them to borrow money at longer terms and to hold less risky kinds of assets. A related rule in the United States, the Volcker Rule, puts further constraints on US banks by restricting which assets they can trade and in what quantity. While these regulations are designed to prevent crises in banking, they may also create unintended costs for financial markets. Specifically, the regulations may unduly constrain the ability of banks to buy and sell securities for their clients, which is a necessary financial service.

The conventional business of a bank is to borrow money at short terms and invest it at long terms. In addition to this, many banks also operate a securities dealer, which is not in the business of investing in the conventional sense. Securities dealers buy and sell securities for distribution to clients, not for investment. Yet Basel III regulates their securities positions as if the dealer were intending to hold its securities to maturity. Thus, some banks protested that Basel III, while appropriate for the risks in conventional banking, has unintended consequences for their business of making markets (see the survey in CGFS 2016 or Duffie 2012). A large empirical literature arose to examine the cost of the regulations, particularly for bond markets, which rely on bank-owned dealers for market making.

While the empirical literature is large and covers many topics, it stands in need of some interpretation, as it has collected findings that can appear to be inconsistent. After the regulations, the literature has found that bid-ask spreads for bonds appear unchanged, and (in some cases) trading volumes have increased. This would appear to be good news. However, in contrast, measures of dealer inventory have fallen, and certain measures of price impact have worsened. These findings are challenging to interpret using past theory on market

making. In most theory, liquidity metrics tend to co-move and get “better” or “worse” all at the same time. However, in the data, the metrics appear to diverge.

In this paper, we offer a theoretical analysis of the regulation that enables a consistent reading of the data. We present a model that can help interpret the data because, in the model, the regulations can have a different effect on different liquidity metrics. The model can deliver the result because it assumes frictions on the balance sheet rather than frictions such as asymmetric information or inventory costs, which are the focus of much past theory. The balance sheet is the right object to study in our context because fixed-income market makers are heavy users of securities financing (Huh and Infante 2018), which inflates the balance sheet in the way the regulations intend to prevent. As a result, we can suggest new ways that empirical work could measure the costs of the regulation, and we deepen the literature’s understanding of how bond markets work.

In the model, a Cournot oligopoly of dealers makes markets with two segments of investor. The investors are divided into two types, buyers and sellers, and they must trade with dealers in order to buy and sell. Since there are two market segments, there are two market prices—the price to buyers and the price to sellers, *i.e.*, a bid and an ask. The existence of a bid-ask spread enables the dealers to make profits as market makers. The dealers, being Cournot, choose how much to buy from sellers and sell to buyers, and they make their choices by optimizing against the two types’ demand curves given the other dealers’ choices. To study regulation in fixed income, which limits the balance sheet, we assume the dealers have balance sheets and thus cannot hold negative quantities. As in Brunnermeier and Pedersen (2008), we assume dealers must hold nonnegative cash and have access to a repo market. To this, we add the assumptions that dealers must also hold nonnegative securities (not just nonnegative cash) and have access to a reverse repo (or specific repo) market.

We impose on the dealers three stylized regulations: on capital, on funding liquidity, and on position. The capital regulation is a cap on the debt-to-equity ratio taken directly from the Basel III Leverage Ratio. The funding-liquidity regulation is more stylized and has two

parts. First, we require dealers to possess liquid, low-risk bonds in a proportion to their short-term borrowing, as in Basel’s Liquidity Coverage Ratio. Second, we require dealers to adjust their risky investments to remain in a proportion with their stable debt (as opposed to short-term debt), as in Basel’s Net Stable Funding Ratio. For more information on these components of Basel III, see our discussion of recent regulation below. Last, as a corollary of the results on capital regulation, we study a simple position limit as a cap on the absolute size of positions, drawn from the US Volcker Rule.

Our main result is that the regulations motivate dealers endogenously to use an agency basis of market making. In agency market making, a securities dealer matches an investor who wants to trade with someone else who will agree to take the other side. This contrasts with the usual basis of market making, the principal basis, in which a dealer itself takes the other side of the trade and holds the position on its balance sheet. Dealers who trade on a principal basis use securities financing to procure cash or securities as needed (Fontaine et al. 2016). This is “balance-sheet intensive,” as one trade has two effects on the balance sheet: one for the trade itself and one for the securities-financing contract. In contrast, the agency model makes less use of the balance sheet, as trades “net out” immediately and never enter the accounts. While agency market making avoids using the balance sheet, it is less flexible than principal market making, because the dealer now depends on its clients to take offsetting trades. Typically, the dealer must motivate its clients to do so by conceding to one of them on price (Choi and Huh 2018).

We find all the regulations encourage agency-based trading but on positions of different sizes. Liquidity regulations add marginal costs and move smaller net positions to agency. Dealers prefer to trade on agency to avoid paying the added marginal cost of buying low-risk bonds or of using more stable debt, as required by liquidity regulation. In contrast, leverage regulation (as well as the Volcker position limit) moves larger net positions to an agency basis. A leverage limit creates a cap on positions, meaning it amounts to a position limit

such as the Volcker Rule, so dealers have no choice but to trade on agency once their position size is sufficiently large.

The outcome of a move to agency is diminished liquidity. The price concession required to motivate investors to trade with one another adds price impact that did not exist before. However, while we find the regulations reduce liquidity via greater price impacts, we find no effect on bid-ask spreads and ambiguous effects on trading volume. The regulations have no effect on the bid-ask spread in our model because the spread is determined by competition among dealers and not by the balance sheet. The regulations also have no necessary effect on trading volume because regulation has countervailing effect on volumes—fewer investors trade because of the weaker liquidity, but agency-based market makers trade two times where principal market makers only needed to trade once, so volume could either increase or decrease depending on parameters. Together, these results imply that bid-ask spreads and trading volumes are imperfect ways to measure the cost of regulation. The outcome is novel, as liquidity metrics in theoretical microstructure typically co-move. Here, we find different frictions leading to different forms of illiquidity.

This reduction in liquidity implies a reduction in social welfare, which we show in the limited setting of a securities market. We do not evaluate the benefits of the regulation for financial stability, as this would require a wider macrofinancial model that is beyond the scope. Nevertheless, while the regulation reduces liquidity and therefore social welfare in the securities market, we do show it makes the securities dealer less risky. Under all the regulation, dealers take smaller net positions, and accordingly they need not compensate their investors as much for risk. Overall, the regulation achieves a cost-risk tradeoff: Dealers are less risky, but liquidity is more costly.

The model predictions can help explain findings in the empirical work on banking regulation. As in the model, regulation has made U.S. dealers reluctant to commit capital to warehousing inventory.¹ The model predicts this leads to an increase in agency-style trad-

¹Bessembinder et al. (2018); Bao et al. (2018); Schultz (2017); Adrian et al. (2017).

ing, which is found by Choi and Huh (2018). To study the liquidity impact, the empirical literature has focused on issuer-specific events that cause sudden, large and directional trade flows.² These have increased the cost of immediacy via price impact, also predicted by the model.

The empirical literature has not found a similar worsening in other liquidity metrics, such as bid-ask spreads or volumes. Average spread-based measures of trading costs appear unaffected, and trading volumes perhaps have even increased.³⁴ The model can reconcile these results with those on price impact because it predicts the illiquidity is not measurable in spreads and volumes. The model also makes new predictions that could motivate future empirical work on bond markets. The price impacts observed in the model are permanent so long as dealers hold a certain inventory. That is, the model predicts an increase in price pressure, a persistent price deviation from the efficient price caused by constraints in the ability of dealers to hold inventory (Hendershott and Menkveld 2014). Price pressure has not been examined in the context of the new regulation. However, evidence from foreign-exchange markets corroborates this prediction, as Pinnington and Shamloo (2016) and Du et al. (2018) find deviations in FX prices attributable to the Basel III Leverage Ratio.

Our paper is related to a growing literature on dealer intermediation that is more generally studying how dealers respond to costs. An and Zheng (2018) also study agency trading and find that a conflict of interest produces a similar irrelevance result on the bid-ask spread. They also argue the metric is a poor measure of costs at dealers. Li and Li (2018) finds similarly that inventory costs are related to the decision of whether to broker trades. Recently, Saar et al. (2019) has studied the effect of Basel III on dealer liquidity and has a distinct finding that the forced move to agency is good for liquidity. The finding that regulatory cost *increases* welfare derives from a special assumption that banks are incumbents with market power that is defeated by the regulation.

²Bao et al. (2018); Dick-Nielsen and Rossi (2018); Schultz (2017).

³Trebbi and Xiao (2017); Adrian et al. (2017); Anderson and Stulz (2017); Mizrach (2015).

⁴However, Bao et al. (2018) find specific events cause mildly wider bid-ask spreads, which the model can match if there is a withdrawal of a dealer from the market after the event.

A. *Recent prudential banking regulation*

The model we present in this paper is a tool to better understand how new banking regulations affect the business of market making. The regulations we have in mind are primarily those recommended by Basel III, a regulatory framework developed by the Basel Committee on Banking Supervision (BCBS), though we also examine the position regulations in the U.S. Volcker Rule, part of the 2010 U.S. Dodd Frank Act.

The BCBS, a forum for central banks and national banking supervisors, has developed three rounds of regulatory frameworks. The first two, the 1988 Accord (Basel I) and Basel II in 2004, proposed minimum capital requirements for regulated banks and minimum standards on their supervision and reporting. Basel II was intended to be implemented by 2008, but the implementation period was interrupted by the 2007-2008 financial crisis, and the BCBS had to respond. Basel I and II had taken an approach to banking regulation that was relatively microprudential, *i.e.*, more focused on risks to individual banks than on risk to the system. However, the crisis demonstrated that liquidity problems create risk for the system as a whole. A shock to one financial institution can propagate to others because institutions depend on one another for short-term funding. Therefore, the BCBS responded by working on a framework that was more macroprudential in that it focused on stopping the propagation of a shock.

In late 2010, it proposed new standards now known as Basel III. Basel III contains new **liquidity requirements**, which intend to enhance banks' ability to withstand a sudden withdrawal in their funding sources. The liquidity requirements ask banks either to borrow at longer terms to finance their assets or, alternatively, to hold more saleable assets relative to their short-term borrowing. Specifically:

- The **Liquidity Coverage Ratio** requires banks to hold “high-quality” assets in proportion to any borrowing with term 30 days or less. These assets can be sold off to raise funds in case of a sudden withdrawal in funding.

- The **Net Stable Funding Ratio** requires banks to fund assets that mature at various terms less than one year with financing that has at least a matching term. Since the funding is locked in, a stop in the rollover of funding does not necessitate an immediate fire sale of the asset it funded.

Basel III also contains **stronger versions of existing capital requirements and a new capital requirement**, which are intended to increase the buffer of loss-absorbing equity at banks. The framework includes a larger minimum of equity and reserves as a percentage of risk-weighted assets (riskier assets require more equity), which meant a revised **Capital Adequacy Ratio**, and also creates:

- The **Leverage Ratio**, a new and non-risk-weighted equity requirement, which requires banks to maintain a quantity of stock and cash equal to at least 3% (in Basel's formulation) of assets. The U.S. Federal Reserve has implemented a ratio of 6% for large banks. This provides an equity cushion for banks against shocks that are not adequately represented by risk-weighted requirements.

The new requirements are more macroprudential because they intend to make banks less interdependent. Banks depend on one another via interbank lending markets, in which banks extend short-term loans to one another. If one bank suddenly stops lending due to funding problems, the problems can propagate to other banks if they are not prepared. As a preventative, the new liquidity regulation asks either that banks have more liquid assets on hand that they can sell to make up for temporary losses in funding or that banks use funding that simply is not short in term in the first place. The Leverage Ratio, in a similar spirit, guarantees a minimum, unconditional cushion of equity so that banks are less runnable in the event of a liquidity crisis.

The regulations target risks inherent to a bank's lending business, which is essentially a business of borrowing short to lend long. Many banks also own and operate a securities dealer, which is a different business. After Basel III was announced, some securities dealers

and other market participants argued that the rules were more appropriate to the lending business than to the securities business and that they had unintended consequences for the liquidity provided by securities dealers. For example, a securities dealer that must temporarily hold a long position for a client typically finances this position using short-term debt (of which the paradigm example is repo), since the dealer does not really intend to invest in the security but to sell it at first opportunity to another client. The new liquidity regulations, nevertheless, require the dealer to fund the position using debt at a longer term, which is costly. The new leverage regulation can be harsher, since a dealer may cease to be able to intermediate while it is financing a temporary but enormous position for a client, and the explicit, numerical position limit in the Volcker Rule acts similarly.

Still, it is unclear to regulators whether the negative consequences are material or, if they are material, then whether they are truly unintended. First, regulation does not apply directly to the securities dealer but to the whole bank, including all its business units. Using transfer pricing, a bank's treasury might pass the costs of regulation to businesses that are more reliant on short-term funding than the securities dealer. Many banks have a substantial short-term funding base that was chosen because it was slightly cheaper and not because the economics of banking demand it. An empirical academic literature arose to check whether the costs were really being passed to dealers, and it appears at least some are. But, to the extent that the regulation did worsen liquidity materially, it could be this is an intended consequence. Dealers are subject to systemic risk as well, and regulators sometimes argue that the price of liquidity was too low before regulation (CGFS 2016).

II. Model

We write a model of Cournot competition among dealers who may buy and sell an asset with two segments of investors. The dealers choose quantities in Nash equilibrium given other dealers' choices and given the investors' demand curves. Since there are two market segments, there are two market prices—the price to buyers and the price to sellers, *i.e.*, a

bid and ask. The existence of a bid-ask spread enables the dealers to make profits as market makers. Dealers have balance sheets and cannot hold “negative” quantities of anything, so they use repo and reverse repo to procure necessary cash or securities.

A. Goods and timing

There are two periods of the game. In the first period, $t = 1$, dealers choose their quantities to trade. In the second period, $t = 2$, trades settle and payoffs realize.

Traded asset: There exists one traded asset in unbounded supply. The asset pays off in cash, has an expected value v in $t = 1$ and realizes a final value V following the end of trading in $t = 2$. The distribution of the asset’s returns at $t = 2$ are given by the function $\phi_V(V)$.

High-quality liquid asset (HQLA): In addition, there exists one HQLA in unbounded supply. The asset represents a “risk-free” security, such as high-quality government debt, that regulation can ask banks to hold. HQLA has a normalized price of 1 and offers a risk-free return. Buying HQLA costs the bank the opportunity of the next-best risky use of its funding. We note the opportunity cost of holding HQLA as r_O .

Repurchase agreements (repo) and reverse-repurchase agreements (reverse repo): There exist repo and reverse-repo agreements in unlimited quantity. A repo is a general-collateral repo, a cash loan that is collateralized by equivalent value of an asset. A reverse repo is a specific repo Duffie (1996), a loan of the specific traded asset that is collateralized by equivalent value of cash. We assume repos in the model have no haircuts, and we assume that the rate on repo and specific reverse-repo agreements is flat, *i.e.*, invariant to the quantity demanded of repo or reverse repo.

General-collateral repo and specific reverse repo are securities-financing contracts that play a central role in bond market making (Choi and Huh 2018; Fontaine et al. 2016; Garriott and Gray 2016). Dealers use their bond inventory as collateral to finance their trading, often using a bond they have purchased as a collateral to finance the very same purchase. Dealers

also use specific reverse repo to procure bonds for sale to clients, in effect “shorting” them, to give themselves time either to wait for an offsetting trade or to source the bond outright.

We will assume that repo and reverse repo can be costly. Empirically, repurchase agreements often charge a rate equal to the overnight rate, meaning they are “costless” from the standpoint of opportunity cost since funding can usually be procured or deposited at the overnight rate. Accordingly, in markets, the rates on repurchase agreements are often expressed in terms of a spread to the overnight rate, and the interpretation of this spread is that repo at a rate above overnight is costly whereas *reverse* repo at a rate *below* overnight is costly. In our model, we assume the rates at which repo and specific reverse repo are available can be costly in terms of this spread: $r_R \geq 0$ and $r_S \geq 0$ respectively. Effectively, we are assuming there is a bid-ask spread in the repo market that weakly surrounds the overnight rate. While the reader may remove this assumption by setting $r_R = 0$ and $r_S = 0$ with no consequences to the results, we find nonzero repo rates are good for two reasons. First, it is of course more realistic that borrowing something should be costly rather than free. Second, it enables us to show the costs of regulation enter as an addition on the repo or reverse repo rates. Using our model, the costs of regulation can be construed as “more costly repo.”

B. Investors

There are two segments of utility-maximizing investors: buyers and sellers. Buyers and sellers are distinguished because of their differing utility functions and because they are unable to transact with each other and must transact via dealers. The investors are price-sensitive and have quadratic preferences, which generate linear demand curves. For notational ease, we write the utility functions from the perspective of the dealers, so buyer utility is a function of securities sold to them S ,

$$U(S) = (v + l_S - P_S)S - \frac{1}{2}\lambda_S S^2, \tag{1}$$

and sellers' utility is a function of securities bought from them B ,

$$U(B) = -(v - l_B - P_B)B - \frac{1}{2}\lambda_B B^2, \quad (2)$$

where v is the asset value, l_S and l_B are level factors that scale investor demand to buy (l_S) or to sell (l_B), P_S and P_B are the prices to buyers and to sellers, and λ_S and λ_B are elasticity parameters that scale the price elasticity of buyers (λ_S) and sellers (λ_B).

Investors maximize utility by choosing B or S given prices P_B and P_S . Once maximized, the utility functions of these investors correspond to linear inverse-demand curves at which dealers can buy and sell the traded asset:

$$P_B = v - l_B + \lambda_B B, \quad (3)$$

$$P_S = v + l_S - \lambda_S S. \quad (4)$$

To interpret the inverse-demand functions, we refer to the level factors as investor demand to buy (l_S) or investor demand to sell (l_B), and together simply as investor demand. Larger values of investor demand act to raise the level of the investors' inverse-demand curves. Last, for notational ease, we define the total price elasticity Λ , as the total price elasticity appears in many formulae:

Definition XXX: *Let $\Lambda = \lambda_B + \lambda_S$ be the sum of the price elasticity parameters.*

C. Dealers

There exist N identical banks indexed by $i \in 1, 2, \dots, N$. Each bank operates a securities dealer, buying from investor sellers and selling to investor buyers, and the dealer arm acts to maximize firm value F_i at time $t = 1$ given the expected market making profits π_i :

$$F_i = \frac{E[\pi_i(V)]}{1 + r_A}, \quad (5)$$

where r_A is a dealer's return on assets (RoA). In conventional models of asset pricing (such as the CAPM and APT), the returns increase linearly in factor risk. The risk at the securities dealer is the variability of the dealer's market-making returns. Thus, RoA can be written:

$$r_A = \beta \sqrt{\int_V \phi_V(V) (r_A(V) - E[r_A(V)])^2 dV} \quad (6)$$

where β is the bank's exposure to risk at its dealer, $r_A(V)$ is the *realized* return given V and $E[r_A(V)]$ is the expected return. Substituting $r_A(V) = \frac{\pi_i(V) - F_i}{F_i}$ and $E[r_A(V)] = \frac{E[\pi_i(V)] - F_i}{F_i}$, and recalling that a dealer maximizes firm value, then the dealer maximizes:

$$F_i = E[\pi_i(V)] - \beta \sqrt{\int_V \phi_V (\pi_i(V) - E[\pi_i(V)])^2 dV}. \quad (7)$$

The market-making profit function π_i is determined by the industrial organization. Again, the dealers are Cournot competitors in the market for investor buyers and the market for investor sellers, which have the inverse-demand curves given in equations 3 and 4. Dealers maximize firm value by:

- selecting a quantity of the traded asset to buy (b_i) and sell (s_i), cognizant that the quantities it chooses will affect prices P_B and P_S given the quantities other banks choose;
- selecting a quantity of repo or reverse repo (as banks are unable to hold negative cash or assets), so that a bank that takes a long position $b_i > s_i$ finances it via the repo market at cost r_R , and a bank that takes a short position $s_i > b_i$ obtains the bonds from the reverse-repo market at a cost r_S ;
- and selecting a weakly positive quantity of HQLA to buy (H_i) at the normalized price 1, which costs the bank the total opportunity cost of $-H_i r_O$.

A dealer's market-making profits are thus,

$$\pi_i(v) = \underbrace{\left(l_B - \lambda_B \sum b_i \right) b_i + \left(l_S - \lambda_S \sum s_i \right) s_i}_{\text{Revenue from market making}} - \underbrace{H_i r_O}_{\text{Cost of HQLA}} - \underbrace{v(b_i - s_i)r_R \mathbb{1}(b_i > s_i)}_{\text{Cost of repo}} - \underbrace{v(s_i - b_i)r_S \mathbb{1}(s_i > b_i)}_{\text{Cost of reverse repo}} \quad (8)$$

The symmetry of the profit function enables a simplification of the radical in equation (7). If the dealer takes a long position, $\pi_i(V) - E[\pi_i]$ simplifies to $(b_i - s_i)(V - v)$. Equally, if the dealer takes a short position, $\pi_i(V) - E[\pi_i]$ simplifies to $(s_i - b_i)(v - V)$. Since b_i and s_i are chosen independent of the asset's final value, the firm value is

$$F_i = E[\pi_i] - \beta |b_i - s_i| \sqrt{\int_v \phi_v (v - V)^2 dv}. \quad (9)$$

What is left in the radical is simply the standard deviation of the asset value, which we denote $\Phi = \sqrt{\int_v \phi_v (v - V)^2 dv}$. Thus, the dealer selects b_i , s_i and H_i to maximize the objective:

$$F_i = E[\pi_i] - \beta |b_i - s_i| \Phi. \quad (10)$$

The β parameter enables us to vary the size of the securities dealers as risk factors within the model banks. Most banks operate several business units, each of which has its own risk profile and funding needs. Banks distribute funding and manage risk across their various units by using transfer pricing, which is represented in our model via the opportunity costs of funding and via the required return on risk. By varying β , we can study markets in which banks have greater or lesser exposure to the risk of their securities dealer, and hence we can study markets in which regulation has greater or lesser impact on banks. To explore the question further, in the appendix, we study a market in which some banks have greater exposure to regulation than others. We expand the model to a case in which there are two types of dealers, those who are more regulated and those who are less regulated (which covers

the case of some dealers having higher β , since regulatory costs arrive linearly). The results do not change sufficiently for us to include it in the main text.

III. Baseline equilibrium

We begin by studying the case of no regulation as a benchmark. In this equilibrium, the bank-owned dealer has no compliance obligations, so it will not purchase HQLA. The dealer will trade on a principal basis if the imbalance in investor demand is large enough to merit the use of costly securities financing. Otherwise, if investor demand to buy and to sell are close enough to equality, the dealer endogenously matches investors on an agency basis, avoiding the costs of securities financing.

For clarity, we expand the objective function in equation (10) to remove the indicator functions. Conditional on an equilibrium in which the dealer chooses $b_i > s_i$, then the dealer is using repo, and it chooses b_i , s_i and H_i to maximize:

$$F_i = \left(l_B - vr_R - \lambda_B \sum b_i \right) b_i + \left(l_S + vr_R - \lambda_S \sum s_i \right) s_i - H_i r_O - \beta \Phi(b_i - s_i). \quad (11)$$

Conditional on an equilibrium in which the dealer chooses $s_i \geq b_i$, then the dealer is using reverse repo, and it chooses b_i , s_i and H_i to maximize:

$$F_i = \left(l_B + vr_S - \lambda_B \sum b_i \right) b_i + \left(l_S - vr_S - \lambda_S \sum s_i \right) s_i - H_i r_O - \beta \Phi(s_i - b_i). \quad (12)$$

For ease of interpretation, we group certain model parameters into cost factors.

Definition 1: Let $\gamma_+ = vr_R + \beta\Phi$ be the cost of taking a long position, and let $\gamma_- = vr_S + \beta\Phi$ be the cost of taking a short position.

The reason we define these cost factors is to group into two variables all the costs a dealer pays when it takes a position. The γ_+ cost represents the costs of long positions, and the γ_- cost represents the costs of short positions. There are costs because whenever a dealer takes

Table I
Regions of dealer market-making in the baseline case

Investor demand	Parameter region
Short	$\lambda_S l_B - \lambda_B l_S < -\Lambda \gamma_-$
Agency	$-\Lambda \gamma_- \leq \lambda_S l_B - \lambda_B l_S < \Lambda \gamma_+$
Long	$\Lambda \gamma_+ \leq \lambda_S l_B - \lambda_B l_S$

any position, it must (a) finance the position and (b) pay a return on assets due to the risk of the position. For long positions, the dealer finances the position using repo, paying vr_B , and, for the risk, it pays the return on assets $\beta\Phi$. Accordingly, γ_+ is just the sum of these two costs. The other cost factor, γ_- , is defined similarly but for short positions. In this case, the dealer procures the necessary assets also by borrowing them, paying vr_S , and by paying a similar return on assets due to the risk of shorting an asset. Then, γ_- is the sum of those two costs.

With the cost factors defined, we now establish there is an equilibrium:

Proposition 1 (Existence of the baseline equilibrium):

There exists a symmetric Cournot equilibrium such that each bank i selects an identical b_i , s_i and H_i to maximize equation 10.

The key to understanding the equilibrium is to know that a dealer changes its basis of trade depending on whether the revenue of taking a position is greater than the cost. If the revenue of taking a position is sufficient, the dealer trades on a principal basis, meaning it takes a nonzero position and pays the cost factor γ_- or γ_+ corresponding to the sign of the position. However, if instead the cost of the position is greater, the dealer trades on agency, endogenously choosing to buy an asset for every asset it sells and taking no position. In short, depending on the size of the trading revenue relative to the costs of taking a position, the dealer exhibits three types of equilibrium behaviour: going short, going long, or trading on agency.

Table I defines the regions in which the dealers go short, go long, or trade on agency. The regions are defined by the size of the balance of investor demands $l_B - l_S$ relative to the size of the cost variables γ_- and γ_+ (the variables are weighted by the relevant price elasticities). The difference in investor demands plays a role in the definition because the revenue to holding a position scales in the position size, as the model has a Cournot core. So when the position can be large ($l_S > l_B$ by a lot, or $l_B > l_S$ by a lot), it merits paying the cost.

A. Quantities

The quantities traded by dealers have the typical Cournot structure. As in classic Cournot, the quantities bought and sold depend linearly on investor demand and on the “costs of production” (γ_- and γ_+), whereas they depend inversely on the number of competitors and on the price elasticities of investor demand.

$$b_i = \begin{cases} \frac{l_B + \gamma_-}{\lambda_B(N+1)} \\ \frac{l_B + l_S}{\Lambda(N+1)} \\ \frac{l_B - \gamma_+}{\lambda_B(N+1)} \end{cases} \quad s_i = \begin{cases} \frac{l_S - \gamma_-}{\lambda_S(N+1)} & \text{Short} \\ \frac{l_B + l_S}{\Lambda(N+1)} & \text{Agency} \\ \frac{l_S + \gamma_+}{\lambda_S(N+1)} & \text{Long} \end{cases} \quad (13)$$

In fact, the quantities are sufficiently like Cournot that those in the short and long regions are simply the Cournot results for the two markets considered in isolation. In contrast, the agency region has a wrinkle. In the agency region, both investor demand variables enter into both the buy and sell quantities. This is because a dealer who is trading on agency is attempting to balance the demands from both markets. Dealers endogenously choose $b_i = s_i$ in this parameter region.

Equation 14 gives the gross quantity traded overall, $Q = N * b_i + N * s_i$, which in empirical studies is measured as total trading volume:

$$Q = \begin{cases} \frac{1}{\lambda_B(N+1)}l_B + \frac{1}{\lambda_S(N+1)}l_S + \frac{\lambda_S - \lambda_B}{\lambda_B \lambda_S(N+1)}\gamma_- & \text{Short} \\ \frac{2}{\Lambda(N+1)}l_B + \frac{2}{\Lambda(N+1)}l_S & \text{Agency} \\ \frac{1}{\lambda_B(N+1)}l_B + \frac{1}{\lambda_S(N+1)}l_S + \frac{\lambda_B - \lambda_S}{\lambda_B \lambda_S(N+1)}\gamma_+ & \text{Long} \end{cases} \quad (14)$$

Gross quantities increase continuously, monotonically and positively in both the investor demand variables l_B and l_S . The more demand for trading, the more gross trading there is. However, gross quantity does not increase positively in both cost variables, which have different signs in different regions, depending on the price elasticities. Gross quantity decreases in cost when a dealer is taking positions for the segment of investors that is more price sensitive.

B. Prices

The prices in equilibrium are also natural outcomes of a Cournot setting. As with quantities, prices depend linearly on investor demand and on the costs of “production” (γ_- for short positions; γ_+ for long), and they depend inversely on the number of competitors.

$$P_B = \begin{cases} v - \frac{1}{N+1}l_B + \frac{N}{N+1}\gamma_- \\ v - \frac{\Lambda + \lambda_S N}{\Lambda(N+1)}l_B + \frac{\lambda_B N}{\Lambda(N+1)}l_S \\ v - \frac{1}{N+1}l_B - \frac{N}{N+1}\gamma_+ \end{cases} \quad P_S = \begin{cases} v + \frac{1}{N+1}l_S + \frac{N}{N+1}\gamma_- & \text{Short} \\ v + \frac{\Lambda + \lambda_B N}{\Lambda(N+1)}l_S - \frac{\lambda_S N}{\Lambda(N+1)}l_B & \text{Agency} \\ v + \frac{1}{N+1}l_S - \frac{N}{N+1}\gamma_+ & \text{Long} \end{cases} \quad (15)$$

Since the markets are segmented, there is a bid price P_B and an ask price P_S . The spread between bid and ask is the source of trading revenue for dealers in the model. The bid and ask display the expected relationship to investor demands. The ask is linearly and positively related to l_S , meaning that buyer demand raises the buying price; and the bid is linearly and negatively related to l_B , meaning that seller demand lowers the selling price. The costs

of taking a position, γ_- or γ_+ , enter linearly to both the bid and ask by the same factor, meaning that asset prices adjust to pass through the position costs to investors (and they pass through with incidence $\frac{N}{N+1}$).

The essence of our results is on display in the agency region of the prices: Investor demand has a larger price impact when a dealer is trading on agency. This can be seen mathematically by comparing the coefficient on investor demand in the short or long regions $\frac{1}{N+1}$ to the same coefficient in the agency region $\frac{\Lambda+N\lambda_i}{(N+1)\Lambda}$, which is larger. Economically, the reason the price impact is greater is because a dealer trading on agency must acquire needed cash or assets from investors. This moves prices because the dealer has to incentivize investors to trade who would not otherwise trade, meaning the dealer has to concede to them on price. In contrast, a dealer in the long or short regions can use repo markets to borrow cash or asset instead of acquiring them outright. In these regions, the dealer responds to investor demand by increasing supply, as borrowing increases the effective supply of a thing since it can be used by two different people at the same time. This is analogous to the way that fractional-reserve banking increases the money supply. Thus, with securities financing, prices do not have to adjust to demand as much as they do in the agency model. Supply can adjust to demand instead.

The bid-ask spread, which is often used empirically as a liquidity proxy, is invariant to most of the model parameters:

$$P_S - P_B = \begin{cases} \frac{l_B+l_S}{N+1} & \text{Short} \\ \frac{l_B+l_S}{N+1} & \text{Agency} \\ \frac{l_B+l_S}{N+1} & \text{Long} \end{cases} \quad (16)$$

There are no securities-financing costs, and moreover no elasticities of demand, in the bid-ask spread. It is determined solely by competitive frictions (the number of firms) and by the total investor demand $l_B + l_S$. This creates the potential for regulation to alter only certain

dimensions of liquidity, for example the price impact, without changing other dimensions, such as the bid-ask spread. Regulation could affect price impacts without affecting bid-ask spreads if it were to act as an amplifier of securities-financing costs, as we show it does in later sections.

C. Returns and risk

The equilibrium RoA is given by $r_A = \beta|b_i - s_i| \sqrt{\int_v \phi_v (v - V)^2 dv}$:

$$r_A = \begin{cases} \frac{\beta\Phi}{\lambda_S(N+1)}l_S - \frac{\beta\Phi}{\lambda_B(N+1)}l_B - \frac{\Lambda\beta\Phi}{\lambda_B\lambda_S(N+1)}\gamma_- & \text{Short} \\ 0 & \text{Agency} \\ \frac{\beta\Phi}{\lambda_B(N+1)}l_B - \frac{\beta\Phi}{\lambda_S(N+1)}l_S - \frac{\Lambda\beta\Phi}{\lambda_B\lambda_S(N+1)}\gamma_+ & \text{Long} \end{cases} \quad (17)$$

The return on assets is a measure of risk-taking by the dealer arm of banks. Naturally, then, it is increasing when taking long positions in l_S , the buyer demand, and it is increasing when taking short positions in l_B , the seller demand. The risk is decreasing in the costs of bearing positions, γ_+ and γ_- , because the dealer charges more for these positions when the cost is large. This represents a benefit to financial *stability* of the cost to financial *efficiency* of taking a position. Notably, the return on assets is zero in the case of agency trade, since in this case the dealer is taking no risk. This represents a strong financial-stability benefit of agency-based trade.

IV. The case of liquidity regulation

Now, we impose liquidity regulation on the dealers. As a reminder, for liquidity regulation, we impose stylizations of the Basel III Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR). Again, the idea of liquidity regulation in the context of repo is to buffer the bank against a failure of its repo liabilities to roll over, in which case the bank would be suddenly unable to finance its inventory and would have to engage in a fire

sale. The LCR prepares the bank by asking it to hold HQLA that it could sell easily in such a case. The NSFR prepares the bank by asking it to “term out” its financing so that rollover failures are less likely to cause a funding crisis in the first place.

Specifically, the LCR asks institutions to hold sufficient assets deemed high-quality to cover all liabilities due in 30 days or less. Thus, we apply the LCR as an obligation to hold quantities of HQLA in proportion to repo liabilities, which in practice are short in term. Banks comply by shifting funds from their next-best investment into HQLA, meaning banks incur the opportunity cost of HQLA, r_O . As for the NSFR, it asks institutions to hold stable, long-term funding in a proportion to investments with terms greater than six months. We apply the NSFR as an obligation to possess long-term funding in proportion to investments, be they in inventory (long positions) or in reverse repo. We assume banks comply by changing their flexible investments, as it is faster to change these investments than it is to conduct an issuance of a long-term bond, which can be planned for months. Banks therefore shift funds from the next-best investment into HQLA, incurring the opportunity cost of HQLA.

Dealers have ways to adapt to regulation that we do not model. However, the avoidance either results in costs of the same nature as those we do model, or we can parameterize the avoidance. For example, dealers may evade the LCR by extending the terms of their repo beyond 30 days (which is arguably an intended consequence of the LCR). But this would mean paying a higher rate than the short-term rate, so it would lead to costs in the same proportion as forced borrowing to purchase HQLA. These costs can be parameterized by raising or lowering γ_+ . Or, dealers may evade the NSFR’s effect specifically on reverse repo by reducing the terms of their reverse repo below six months. As reverse repo are unlikely to have such long terms in the first place, the weak or avoided effect can be parameterized as γ_- being small or even zero, whereas the impact of the NSFR on bond inventories can remain large (γ_+ large).

Formally, we model the regulation via two assumptions:

Assumption 1 (Liquidity regulation on repo and on long positions): If the bank takes a long position ($b_i > s_i$), it is using repo and holding positive inventory, so it must hold $H_i \geq \alpha_+ v(b_i - s_i)$, where α_+ is proportion of HQLA required by regulation.

Assumption 2 (Liquidity regulation on reverse repo): If the bank takes a short position ($s_i > b_i$), it uses reverse repo, so it must hold $H_i \geq \alpha_- v(s_i - b_i)$, where α_- is the proportion of HQLA required by regulation.

For clarity, we restate the objective function (equation 10) to remove the indicator functions and to impose the liquidity constraints. Conditional on an equilibrium in which the dealer chooses $b_i > s_i$, then the dealer is using repo, and it chooses b_i , s_i and H_i to maximize:

$$F_i = \left(l_B - vr_R - \lambda_B \sum b_i \right) b_i + \left(l_S + vr_R - \lambda_S \sum s_i \right) s_i - H_i r_O - \beta \Phi(b_i - s_i) \quad (18)$$

s.t. $H_i \geq \alpha_+ v(b_i - s_i)$.

Conditional on an equilibrium in which the dealer chooses $b_i > s_i$, then the dealer is using repo, and it chooses b_i , s_i and H_i to maximize:

$$F_i = \left(l_B + vr_S - \lambda_B \sum b_i \right) b_i + \left(l_S - vr_S - \lambda_S \sum s_i \right) s_i - H_i r_O - \beta \Phi(s_i - b_i) \quad (19)$$

s.t. $H_i \geq \alpha_- v(s_i - b_i)$.

For ease of interpretation, we again group certain model parameters into cost factors.

Definition 2: Let $\Gamma_+ = v(r_R + \alpha_+ r_O) + \beta \Phi$ be the cost of taking a long position, and let $\Gamma_- = v(r_S + \alpha_- r_O) + \beta \Phi$ be the cost of taking a short position.

These cost factors are capitalized versions of the cost factors used in the baseline case. We do this to highlight that the cost factors will play the exact same role in the equilibrium objects in the case of liquidity regulation. Moreover, $\Gamma_- > \gamma_-$, and $\Gamma_+ > \gamma_+$. As we will show, this enables us to construe liquidity regulation as an increase in the costs of holding repo and reverse repo.

With the cost factors defined, we now establish there is an equilibrium:

Proposition 2 (Existence of an equilibrium under liquidity regulation):

There exists an equilibrium such that each bank i selects an identical b_i , s_i and H_i to maximize equation (18) or (19), depending on whether the banks take a long ($b_i \geq s_i$) or short ($s_i \geq b_i$) position in equilibrium.

The key to understanding how liquidity regulation changes the equilibrium is to see that the regulation enlarges the region of investor demands for which dealers trade on agency.

Proposition 3 (Agency Trading with Liquidity Regulation):

The region of balances of investor demand $l_B - l_S$ in which market makers take zero position is larger when there is liquidity regulation relative to the baseline.

Table II defines the regions in which the dealers go short, go long, or trade on agency under liquidity regulation. As in the baseline case, the regions are defined by the size of the balance in investor demands $l_B - l_S$ relative to the size of the costs of taking a position. However, these costs are now written as upper-case Γ_- and Γ_+ rather than lower case γ_- and γ_+ . As the new cost factors are larger than the cost factors in the baseline case, the agency region spans a larger interval of $l_B - l_S$.

A. Quantities

We will not report many equations for the case of liquidity regulation because all the results from the baseline case follow for this case with the exact same formulae—except for a substitution of the new cost factor Γ_- for the old γ_- and a substitution of the new cost factor Γ_+ for the old γ_+ . To illustrate the results do not change save for this substitution, we

Table II
Regions of dealer market making in the case of liquidity regulation

Investor demand	Parameter region
Constrained short	$\lambda_S l_B - \lambda_B l_S < -\Lambda \Gamma_-$
Constrained agency	$-\Lambda \Gamma_- \leq \lambda_S l_B - \lambda_B l_S < \Lambda \Gamma_+$
Constrained long	$\Lambda \Gamma_+ \leq \lambda_S l_B - \lambda_B l_S$

repeat the formulae for the equilibrium quantities bought and sold, b_i and s_i . However, to economize, we remove the other equations, since they are the same except for the substitution of the new cost factors.

$$b_i = \begin{cases} \frac{l_B + \Gamma_-}{\lambda_B(N+1)} \\ \frac{l_B + l_S}{\Lambda(N+1)} \\ \frac{l_B - \Gamma_+}{\lambda_B(N+1)} \end{cases} \quad s_i = \begin{cases} \frac{l_S - \Gamma_-}{\lambda_S(N+1)} & \text{Constrained short} \\ \frac{l_B + l_S}{\Lambda(N+1)} & \text{Constrained agency} \\ \frac{l_S + \Gamma_+}{\lambda_S(N+1)} & \text{Constrained long} \end{cases} \quad (20)$$

Proposition 4 (Dealer positions under liquidity regulation):

For given asset demands, the net position ($|b_i - s_i|$) taken by banks with liquidity regulation is weakly less than the net position taken by banks in the baseline.

Since the costs of taking a position are larger, naturally the size of the positions taken by dealers is smaller. It is smaller in the long and short regions and weakly smaller in the agency region (as the dealer already took a zero position in the baseline case). The prediction that regulation pushes dealers to hold smaller positions is consonant with results in the empirical literature (Bessembinder et al. 2018; Bao et al. 2018; Schultz 2017; Adrian et al. 2017).

While the model yields a monotonic prediction for dealer inventory, it does not for the trading volume. There is an ambiguous effect on the gross quantity traded, given by $Q = N * b_i + N * s_i$. The direction of the effect depends on the price elasticities.

Proposition 5 (Trading volume under liquidity regulation):

(i) Compared to the baseline where banks would have taken a long position ($\Lambda\gamma_+ \leq \lambda_S l_B - \lambda_B l_S$), when banks are subject to liquidity regulation, gross quantity traded increases if $\lambda_B > \lambda_S$, while gross quantity traded decreases if $\lambda_S > \lambda_B$.

(ii) Compared to the baseline where banks would have taken a short position ($\lambda_S l_B - \lambda_B l_S < -\Lambda\gamma_-$), when banks are subject to liquidity regulation, gross quantity traded increases if $\lambda_S > \lambda_B$, while gross quantity traded decreases if $\lambda_B > \lambda_S$.

Regulation has ambiguous effects on trading volume even though it decreases positions at dealers because an agency dealer trades twice for every position—once to take the position, as before, but once again to shift the position to another investor. So, trading volume can *increase* after liquidity regulation if the dealer compensates by trading much more with investors to procure needed cash or assets. The dealer does so to reduce its use of costly securities financing.

Specifically, if dealers are taking long positions, and if buyers have lower elasticity than sellers, then trading volume will be higher. In this case, the dealer can easily offload inventory acquired from sellers to buyers at a low price impact. So, the dealer trades more on agency, raising trading volume. However, if sellers have the lower elasticity, then this is too expensive. The dealer instead simply raises prices to clear markets, resulting in less trading volume.

The ambiguous result calls into question the usefulness of measuring trading volume and of using other volume metrics (such as days of zero trades) to assess the costs of regulation, which forms part of the analysis in Trebbi and Xiao (2017) or Adrian et al. (2017). This result is consonant with findings that dealers are moving toward a relatively greater agency basis of trade (Bessembinder et al. 2018; Choi and Huh 2018).

B. Prices

Under liquidity regulation, the same investor demand can cause a larger price impact. The price impact increases for two reasons: First, the position costs Γ_- and Γ_+ are greater, so dealers have larger costs to pass on to investors than they did in the baseline case. Second, the regulation enlarges the region of investor demands $l_B - l_S$ for which the dealer trades on agency. As discussed in the baseline case, the price impact of investor demand to buy or to sell is greatest in the agency region, so the greater price impact of agency-based dealing now affects a wider range of investor demands. The implication of this result is that a good way to assess the costs of regulation is by looking at price impacts, particularly when there are large imbalances of investor demand to buy or to sell. Empirical work in Dick-Nielsen

and Rossi (2018), Bao et al. (2018) and Schultz (2017) has found increases in price impacts, particularly when immediacy is demanded.

However, while price impacts are higher under liquidity regulation, we find the bid-ask spread is unchanged. It was never a function of the cost of taking a position, so the formula is unaffected by liquidity regulation. Economically, this happens because, in an imperfectly competitive market, the spread between bid and ask derives from competitive frictions. But it has no relation to balance-sheet capacity, since dealers simply raise or lower *both* their bid and ask to express the scarcity of the balance sheet. In other words, liquidity regulation has an asset-pricing effect but not an effect on marginal transactions costs.

Proposition 6 (Bid-ask spreads under liquidity regulation):

There is no change in bid-asks spreads when banks are subject to liquidity regulation.

The neutrality result on the bid-ask spread, like the ambiguity result on trading volume, calls into question the usefulness of measuring bid-ask spreads and similar metrics (such as effective spreads or the Roll metric) to assess the costs of regulation, which forms part of the analysis in Bao et al. (2018), Trebbi and Xiao (2017) and Anderson and Stulz (2017). The bid-ask spread can remain *unchanged* after regulation if there is no change in the competitiveness of the dealer sector. Empirically, the data show an increase in competition from nonregulated dealers (Bao et al. 2018), which would help explain findings of decreased bid-ask spreads in certain contexts (Bao et al. 2018; Trebbi and Xiao 2017).

C. Returns and risk

Since the banks take smaller positions and charge more for them, naturally the banks take less risk and deliver a lower return. This is an intended consequence of the regulation in the sense that it makes bank-owned dealers safer.

Proposition 7 (Return on bank assets under liquidity regulation):

Since $\Gamma_+ > \gamma_+$ and $\Gamma_- > \gamma_-$, the equilibrium return on bank assets is weakly lower under

liquidity regulation than in the baseline. That is to say, that the standard deviation of banks' returns is lower under liquidity regulation.

D. Welfare

Last, we consider the welfare effects of liquidity regulation. As we mentioned in the introduction, we do not attempt to write a full model of the macroeconomy in which greater financial stability could improve social welfare. Thus, the welfare effects that we study should be interpreted as the effect on just the securities market and not on macrofinancial stability. Nevertheless, in assessing the costs of regulation, these are important costs to include in any accounting of the costs and benefits.

Ex ante welfare is measured by the total expected surplus of sellers ($\frac{1}{2} * N * (P_B - v + l_B) * b_i$), buyers ($\frac{1}{2} * N * (v + l_B - P_S) * s_i$), and banks ($N * (v - P_B) * b_i + N * (P_S - v) * s_i$). For the purpose of the computation, we consider the bank's cost for the asset to be v . This cost does not include costs due to repo (r_R or r_S), regulation (r_O), or risk (r_A). We assume that these costs are passed efficiently to other, unmodelled agents and thus that these costs do not create deadweight losses.

Proposition 8 (Welfare after liquidity regulation):

(i) Compared to the baseline where banks would have taken a long position ($\Lambda\gamma_+ \leq \lambda_S l_B - \lambda_B l_S$), when banks are subject to liquidity regulation, welfare declines.

(ii) Compared to the baseline where banks would have taken a short position ($\lambda_S l_B - \lambda_B l_S < -\Lambda\gamma_-$), when banks are subject to liquidity regulation, welfare declines.

Welfare declines because the price impact of trading is higher, and so investors who previously traded do so in smaller quantities or not at all. Price impacts go up because of the change in the basis of market making, which is at the core of our results. Principal-based market making can support the market more efficiently because dealers can borrow necessary cash or assets to meet investor demand. The dealer is meeting demand by adjusting the effective supply of cash or assets, as borrowing increases the effective supply of a thing

since it can be used by two different people at the same time. This is analogous to the way that fractional-reserve banking increases the money supply. Thus, with securities financing, prices do not have to adjust to demand as much as they do in the agency model. Supply can adjust to demand instead.

V. The case of capital regulation

Now, we remove liquidity regulation and instead impose capital regulation on the banks. As a reminder, for capital regulation, we impose a stylization of the Basel III Leverage Ratio (LR), which is a cap on the ratio of a bank’s debt to its equity. Again, the idea of capital regulation in the context of repo is to prevent the bank from borrowing too large a sum of money relative to its equity buffer, which may not be large enough to cover a sudden inability to obtain secured funding.

We do not treat the case of the Capital Adequacy Ratio in this paper, as the Capital Adequacy Ratio was an existing regulation, whereas the Leverage Ratio is net new. The Capital Adequacy Ratio asks for a risk-weighted equity buffer (with different weights applied to different assets), so it cannot be addressed directly under the current model. A single constraint on minimum equity is a likely direction of future regulation since it is arguably a stronger regulation (Greenwood et al. 2017). For example, in May 2019, the Financial Stability Board announced a study group that is looking at unconditional loss absorbency for systemically important banks.

To limit banks’ leverage, we assume they have an initial debt and an initial equity. We assume the bank has allocated some fraction of its capital structure E to equity. When a bank enters into a long position and initiates a repo (but not a reverse-repo) contract, it increases its level of debt by $v(b_i - s_i)$. We impose a leverage constraint such that the bank’s leverage can be no higher than some fraction ψ :

$$\frac{v(b_i - s_i)}{v(b_i - s_i) + E} \leq \psi, \tag{21}$$

which is equivalent to:

$$b_i \leq s_i + \frac{\psi E}{v(1-\psi)}. \quad (22)$$

For simplicity, we denote $\Psi = \frac{\psi E}{v(1-\psi)}$. Thus, the firm's optimization problem if it wishes to take a long position is to select b_i , s_i and H_i to maximize:

$$\begin{aligned} F_i &= (l_B - vr_R - \lambda_B \sum b_i)b_i + (l_S + vr_R - \lambda_S \sum s_i)s_i - H_i r_O - \beta |b_i - s_i| \Phi. \\ \text{s.t. } \quad b_i &\leq s_i + \Psi \end{aligned} \quad (23)$$

The firm's optimization problem if it wishes to take a short position is identical to the baseline case. Reverse repo is not accounted as debt under accounting principles used at banks.

Proposition 9 (Existence of an equilibrium under capital regulation):

There exists an equilibrium such that each bank i selects an identical b_i , s_i and H_i to maximize either Equation 23 or 12, depending on whether the banks take a long ($b_i \geq s_i$) or short ($s_i \geq b_i$) position in equilibrium.

The key to understanding how capital regulation changes a dealer's behaviour is to see that it creates a region of investor demand in which dealers trade the *marginal* asset on agency. Table III defines the regions in which the dealers, as before, go short, go long, or trade on agency. However, in contrast to the baseline case and the case of liquidity regulation, there is now a fourth region. What distinguishes this region is that the dealer holds a positive inventory, so it is not trading *purely* on agency, yet the dealer is constrained not to increase its inventory. Therefore, any additional asset that it buys in excess of the constrained inventory is an asset that must also be sold. We describe this as trading the marginal asset on agency.

Table III
Regions of dealer market-making under capital regulation

Investor demand	Parameter region
Short	$\lambda_S l_B - \lambda_B l_S < -\Lambda \gamma_-$
Agency	$-\Lambda \gamma_- \leq \lambda_S l_B - \lambda_B l_S < \Lambda \gamma_+$
Long	$\Lambda \gamma_+ \leq \lambda_S l_B - \lambda_B l_S < \Lambda \gamma_+ + \lambda_B \lambda_S (N + 1) \Psi$
Marginal agency	$\Lambda \gamma_+ + \lambda_B \lambda_S (N + 1) \Psi \leq \lambda_S l_B - \lambda_B l_S$

A. Quantities

The parameter regions of short, agency, and long are preserved from the baseline case. The regions have the same boundaries, and the dealer buys and sells the same quantities in each of the regions. In the new parameter region, marginal agency, the dealer buys and sells the same quantities as in the agency case plus a net position Ψ , which is the maximum inventory it can carry under the leverage constraint.

$$b_i = \begin{cases} \frac{l_B + \gamma_-}{\lambda_B (N + 1)} \\ \frac{l_B + l_S}{\Lambda (N + 1)} \\ \frac{l_B - \gamma_+}{\lambda_B (N + 1)} \\ \frac{l_B + l_S}{\Lambda (N + 1)} + \frac{\lambda_S \Psi}{\Lambda} \end{cases} \quad s_i = \begin{cases} \frac{l_S - \gamma_-}{\lambda_S (N + 1)} & \text{Short} \\ \frac{l_B + l_S}{\Lambda (N + 1)} & \text{Agency} \\ \frac{l_S + \gamma_+}{\lambda_S (N + 1)} & \text{Long} \\ \frac{l_B + l_S}{\Lambda (N + 1)} - \frac{\lambda_B \Psi}{\Lambda} & \text{Marginal agency} \end{cases} \quad (24)$$

The new region, marginal agency, is the region of interest for empirical study. In this region, the dealer's position is smaller than in the baseline case, which is natural as its position is constrained.

Proposition 10 (Net positions under capital regulation):

For given asset demands, the net position ($|b_i - s_i|$) taken by banks with capital regulation is weakly less than the net position taken by banks in the baseline if they become leverage constrained ($\Lambda \gamma_+ + \lambda_B \lambda_S (N + 1) \Psi \leq \lambda_S l_B - \lambda_B l_S$). When banks are not leverage constrained, capital regulation has no impact on net positions.

However, as was seen in the case of liquidity regulation, the prediction on trading volume is ambiguous. Gross quantity traded, given by $Q = N * b_i + N * s_i$, can increase or decrease in equilibrium depending on the elasticities.

Proposition 11 (Gross quantity under capital regulation):

(i) *Compared to the baseline, gross quantity traded is unchanged with capital regulation unless banks are leverage constrained.*

(ii) *When banks become leverage constrained, gross quantity traded increases if $\lambda_B > \lambda_S$ while gross quantity traded decreases if $\lambda_S > \lambda_B$.*

Capital regulation has ambiguous effects on trading volume for the same reasons as does liquidity regulation. Trading volume can *increase* under capital regulation if the dealer compensates by trading much more with investors to procure needed cash or assets. In the leverage-constrained region, if buyers have lower elasticity than sellers, then the dealer can procure cash from buyers for purchasing from sellers at a low price impact. So, the dealer trades more on agency, resulting in more trading volume. As before, the ambiguous result calls into question the usefulness of measuring trading volume and of using other volume metrics (such as days of zero trades) to assess the costs of regulation. This result is again in agreement with findings that dealers are moving toward a relatively greater agency basis of trade.

B. Prices

Under capital regulation, as under liquidity regulation, the same investor demand can cause a larger price impact. Capital regulation creates a new region of investor demands (marginal agency) for which the dealer trades on agency. As discussed in the baseline case, the price impact of investor demand to buy or to sell is greatest in the agency region. Under capital regulation, the greater price impact of agency-based dealing now affects a wider range of investor demands. The implication of this result is, again, that a good way to assess the costs of regulation is by looking at price impacts, particularly when there are large imbalances

of investor demand to buy or to sell, as in Dick-Nielsen and Rossi (2018), Bao et al. (2018) and Schultz (2017).

$$P_B = \begin{cases} v - \frac{1}{N+1}l_B + \frac{N}{N+1}\gamma_- \\ v - \frac{\Lambda + \lambda_S N}{\Lambda(N+1)}l_B + \frac{\lambda_B N}{\Lambda(N+1)}l_S \\ v - \frac{1}{N+1}l_B - \frac{N}{N+1}\gamma_+ \\ v - \frac{\Lambda + \lambda_S N}{\Lambda(N+1)}l_B + \frac{\lambda_B N}{\Lambda(N+1)}l_S \\ + \frac{\lambda_B \lambda_S N}{\Lambda}\Psi \end{cases} \quad P_S = \begin{cases} v + \frac{1}{N+1}l_S + \frac{N}{N+1}\gamma_- & \text{Short} \\ v + \frac{\Lambda + \lambda_B N}{\Lambda(N+1)}l_S - \frac{\lambda_S N}{\Lambda(N+1)}l_B & \text{Agency} \\ v + \frac{1}{N+1}l_S - \frac{N}{N+1}\gamma_+ & \text{Long} \\ v + \frac{\Lambda + \lambda_B N}{\Lambda(N+1)}l_S - \frac{\lambda_S N}{\Lambda(N+1)}l_B & \text{Marginal agency} \\ + \frac{\lambda_B \lambda_S N}{\Lambda}\Psi \end{cases} \quad (25)$$

The prices in the marginal-agency region are the same as in the agency region, except for a factor of Ψ . The reason prices are otherwise the same as in the agency region is that the dealer handles the marginal purchase on an agency basis, so prices in the marginal-agency region move with investor demand the same way that prices do in the agency region. The factor of Ψ is there because prices in the agency region bear a spread on top of the agency prices due to the cost of taking the position permitted by the leverage cap.

Although capital regulation creates a new region of agency trading, which makes price impact worse in that region, there is again no affect on the bid-ask spread.

Proposition 12 (Bid-ask spreads under capital regulation):

There is no change in bid-asks spreads when banks are subject to leverage regulation.

The neutrality result on the bid-ask spread, like the ambiguity result on trading volume, again calls into question the usefulness of measuring bid-ask spreads and similar metrics (such as effective spreads or the Roll metric) to assess the costs of regulation. Empirically, as mentioned, the data show some increase in the presence of nonregulated dealers, which could explain findings of decreased bid-ask spreads.

C. Returns

Capital regulation limits the banks only in one region of investor demands. In this region, the banks take smaller positions, so naturally they take less risk and deliver a lower return. This is an intended consequence of the regulation in the sense that it makes bank-owned dealers safer.

Proposition 13 (Return on bank assets under capital regulation):

When banks become leverage constrained ($\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi \leq \lambda_S l_B - \lambda_B l_S$), their return on assets is lower under leverage regulation. When banks are not leverage constrained, leverage regulation has no impact on bank asset returns.

D. Welfare

Last, we consider the welfare effects of leverage regulation. As in the previous section on welfare, *ex ante* welfare is measured by the total expected surplus of sellers, buyers, and banks; we assume the costs due to repo, regulation and risk are efficiently passed to other, unmodelled agents and so do not create deadweight losses.

Proposition 14 (Welfare under capital regulation):

(i) Compared to the baseline, welfare is unchanged with leverage regulation unless banks are leverage constrained.

(ii) When banks become leverage constrained, welfare declines.

Welfare declines under capital regulation for the same reason that it does under liquidity regulation. The leverage cap increases the price impact of trading when investor demand is sufficiently high to make the dealers leverage-constrained, and dealers trade the marginal asset on agency when they are constrained. Fewer investors who would have traded do so after regulation, and those who do trade do so at smaller quantities and worse prices. Again, price impacts go up because of the change in the basis of market making, which is at the core of our results. Principal-based market making can support the market more efficiently

because dealers can borrow necessary cash or assets to meet investor demand, whereas under agency-based market making, dealers must rely on clients to meet investor demand.

E. Analysis of the Volcker Rule as a corollary of capital regulation

The optimization problem in equation (23) rewrites the cap on leverage as a cap on the size of positive positions: $b_i - s_i \leq \Psi$. This constraint is isomorphic to a position limit, such as that introduced in section 7 of the US Dodd-Frank Act (see Bao et al. 2018 for a more complete analysis of the way the Volcker Rule creates a position limit). Thus, at least for positive positions, the analysis of the Volcker Rule is exactly the same as that for capital regulation. For negative positions, the equal but opposite constraint $s_i - b_i \leq \Psi$ produces the same results for a new, opposite-signed region of investor demands.

VI. Conclusions

The model we present in this paper is a tool to better understand how recent policy affects the business of market making. Our summary finding is that regulations such as those proposed in Basel III (and, by corollary, the Volcker Rule) are motivating a shift by securities dealers to an agency basis of market making. Since agency market making is a less flexible way to make markets, investors pay higher costs for liquidity via greater price impacts, but, in return, securities dealers are also less risky.

The model can match findings from the empirical literature, including by reconciling facts that might be seen as being inconsistent. Specifically, the empirical literature has documented minor impacts to bid-ask spreads, including even benefits in certain cases, and it has found increases in trading volumes. These could be interpreted as good news for the costs of regulation. Yet the literature has also found negative effects on dealer inventories and price impacts, which could be interpreted as bad news. We show the findings of the empirical literature have a consistent interpretation because they are the measurable consequences of a move to agency-based market making. The move makes balance-sheet space more scarce,

which worsens price impacts as these are related to the ability of dealers to take on inventory, but it does not have to affect bid-ask spreads because these are generated by a completely separate friction.

We argue that the model's ability to match the data means that the model can provide reliable insights to policymakers. Since the model is primarily a tool to understand policy, we enumerate the implications of our findings.

1. The cost-risk tradeoff effected by the regulation is an intended consequence.

Balance-sheet regulation is effective at reducing risk at dealers *because* it imposes costs for financial markets and not in spite of the costs. In other words, the illiquidity is not an “unintended consequence.” While the model does not permit a judgement about whether the benefits to system-wide risk merit paying the welfare costs, we can say that the costs are part of the way the regulations work. The regulation adds cost to the financing exposures taken by banks that rise in proportion to the exposures additively with the repo rate. In other words, the regulations make it more costly to take on risk, and it is not surprising that this cost is passed to investor clients who, in part, are motivating the dealer to take on risk. While the costs of regulation might be too high, it could also be that liquidity was underpriced relative to externalities before the financial crisis.

2. The costs of balance-sheet regulation can best be assessed using measures directly related to balance-sheet costs.

Most studies of the costs of Basel III or Volcker are limited by data and cannot observe the funding costs at securities dealers or the transactions decisions at securities investors (including decisions not to invest). Instead, studies use proxies for the costs, and a common proxy for liquidity costs is the bid-ask spread (itself proxied by, for example, the Roll measure or by advertised spreads). The model advice is that policy analysts should proxy the costs using asset-pricing effects rather than using the usual liquidity proxies such as changes to

marginal transactions costs or to transactions volumes. Analysts could look to measure price premia after large flows of trade or during high market volatility, as does current work. Still more directly, they could move to quantifying the overall premium on inventory, “price pressure,” in corporate-bond markets (Hendershott and Menkveld 2014). Friewald and Nagler (2016) gives a methodology to measure price pressure that is useful for fixed income.⁵ The methodology has been used to study the cross-section of price pressure, but it has not yet been used directly to assess the costs of regulation in the time series.

3. The costs of balance-sheet regulation can be attributed among the regulations by studying the cross-section of the costs by dealer position sizes.

Regulatory costs can be attributed to particular regulations based on the sizes of positions taken by dealers. This is because different regulations move different position sizes to agency market making. Liquidity regulation enlarges the existing region of agency trade in the baseline case, which was for small positions only. It pushes dealers who were on the margin between principal and agency market making over to the agency basis. So a finding of greater illiquidity for small trades or from dealers who hold small inventories relative to risk can be attributed to liquidity regulation. However, leverage regulation (and the position limits of Volcker) move large positions to an agency basis. In this case, a dealer runs into a hard constraint on their ability to take on a position beyond some size. Thus, a finding of greater illiquidity for large trades or from dealers who hold large inventories relative to risk can be attributed moreso to leverage regulation or Volcker.

4. The choice of regulatory tool depends on whether the regulator is more concerned about the price of liquidity or about the risk of large positions.

As a corollary of 3, regulators can make a choice between stronger liquidity or stronger capital (or position) regulation by considering the type of financial-market vulnerability they wish to address. Liquidity regulation adds costs to taking positions regardless of the position

⁵Parts of Friewald and Nagler (2016) are now covered by Friewald and Nagler (2019).

size, so if the regulator sees the vulnerability as one of underpriced risk, liquidity regulation may be the best tool. However, if the regulator sees the vulnerability as one of excessively large exposures due to large positions or large quantities of leverage, then capital regulation (which does not add costs to most position sizes) is the more direct tool.

As we have noted, capital regulation and position regulation have similar action in our model, so the decision of which to implement must come down to their other effects that we do not model. A chief difference between Volcker and the Leverage Ratio is that a bank could (and might optimally) adjust to Volcker by financing its securities dealer on a greater debt basis, reducing the equity buffer, since under Volcker the dealer makes fewer and safer returns. In contrast, this is precisely the scenario that the Leverage Ratio rules out. For this and for other reasons, a limit on debt financing may be preferable to Volcker.

REFERENCES

- Adrian, Tobias, Michael Fleming, Or Shachar, and Erik Vogt, 2017, Market liquidity after the financial crisis, *Annual Review of Financial Economics* 9, 43–83.
- An, Yu, and Zeyu Zheng, 2018, Conflicted immediacy provision, *Available at SSRN 2868280* .
- Anderson, Mike, and René M Stulz, 2017, Is post-crisis bond liquidity lower?, Technical report, National Bureau of Economic Research.
- Bao, Jack, Maureen O’Hara, and Xing Alex Zhou, 2018, The Volcker Rule and corporate bond market-making in times of stress, *Journal of Financial Economics* .
- Bessembinder, Hendrik, Stacey E Jacobsen, William F Maxwell, and Kumar Venkataraman, 2018, Capital commitment and illiquidity in corporate bonds, *Journal of Finance* .
- Brunnermeier, Markus K, and Lasse Heje Pedersen, 2008, Market liquidity and funding liquidity, *The Review of Financial Studies* 22, 2201–2238.
- Choi, Jaewon, and Yesol Huh, 2018, Customer liquidity provision in corporate bond markets, *Available at SSRN: <https://ssrn.com/abstract=2848344>* .
- Committee on the Global Financial System, 2016, Fixed income market liquidity, *CGFS Papers* 55.
- Dick-Nielsen, Jens, and Marco Rossi, 2018, The cost of immediacy for corporate bonds, *Available at SSRN: <https://ssrn.com/abstract=2139403>* .
- Du, Wenxin, Alexander Tepper, and Adrien Verdelhan, 2018, Deviations from covered interest rate parity, *Journal of Finance* 73, 915–957.
- Duffie, Darrell, 1996, Special repo rates, *The Journal of Finance* 51, 493–526.

- Duffie, Darrell, 2012, Market making under the proposed Volcker Rule, *Available at SSRN: <https://ssrn.com/abstract=1990472>* .
- Fontaine, Jean-Sébastien, Corey Garriott, and Kyle Gray, 2016, Securities financing and bond market liquidity, *Financial System Review* .
- Friewald, Nils, and Florian Nagler, 2016, Dealer inventory and the cross-section of corporate bond returns, *Available at SSRN: <https://ssrn.com/abstract=2526291>* .
- Friewald, Nils, and Florian Nagler, 2019, Over-the-counter market frictions and yield spread changes, *Journal of Finance* .
- Garriott, Corey, and Kyle Gray, 2016, Canadian repo market ecology, *Bank of Canada Staff Discussion Paper* .
- Greenwood, Robin, Jeremy C Stein, Samuel G Hanson, and Adi Sunderam, 2017, Strengthening and streamlining bank capital regulation, *Brookings Papers on Economic Activity* 2017, 479–565.
- Hendershott, Terrence, and Albert J Menkveld, 2014, Price pressures, *Journal of Financial Economics* 114, 405–423.
- Huh, Yesol, and Sebastian Infante, 2018, Bond market liquidity and the role of repo, *Available at SSRN: <https://ssrn.com/abstract=2773678>* .
- Li, Jiacui, and Wenhao Li, 2018, Brokers or dealers? trading intermediation across markets and over time, *Available at SSRN 2760457* .
- Mizrach, Bruce, 2015, Analysis of corporate bond liquidity, *Research Note, FINRA Office of the Chief Economist* .
- Pinnington, James, and Maral Shamloo, 2016, Limits to arbitrage and deviations from covered interest rate parity, *Bank of Canada Working Paper* .

Saar, Gideon, Jian Sun, Ron Yang, and Haoxiang Zhu, 2019, From market making to match-making: Does bank regulation harm market liquidity?, *Available at SSRN 3399063* .

Schultz, Paul, 2017, Inventory management by corporate bond dealers, *Available at SSRN: <https://ssrn.com/abstract=2966919>* .

Trebbi, Francesco, and Kairong Xiao, 2017, Regulation and market liquidity, *Management Science* .

A Appendix: Proofs

A. Proof of Proposition 1

The firm selects b_i , s_i and H_i to maximize:

$$F_i = E[\pi_i] - \beta|b_i - s_i|\Phi. \quad (26)$$

First, consider an bank which wishes to take a long position ($b_i > s_i$). The firm selects b_i , s_i and H_i to maximize Equation 11. We take first-order conditions and subsequently impose symmetry, such that each bank selects an identical b_i , s_i and H_i . The resulting solutions are $b_i = \frac{l_B - \gamma_+}{\lambda_B(N+1)}$, $s_i = \frac{l_S + \gamma_+}{\lambda_S(N+1)}$ and $H_i = 0$.

By inspection, the bank will optimally select a long position ($b_i > s_i$) when $\lambda_S l_B - \lambda_B l_S > \Lambda \gamma_+$.

Second, consider an bank which wishes to take a short position ($s_i > b_i$). The firm selects b_i , s_i and H_i to maximize Equation 12. We take first-order conditions and subsequently impose symmetry, such that each bank selects an identical b_i , s_i and H_i . The resulting solutions are $b_i = \frac{l_B + \gamma_-}{\lambda_B(N+1)}$, $s_i = \frac{l_S - \gamma_-}{\lambda_S(N+1)}$ and $H_i = 0$.

By inspection, the bank will optimally select a short position ($s_i > b_i$) when $\lambda_B l_S - \lambda_S l_B < -\Lambda \gamma_-$. This region is mutually exclusive over the region under which a firm wishes to take a long position, and thus there are no incentive compatibility issues.

Finally, consider a bank which does not want to take either a long or short position ($-\Lambda \gamma_- \leq \lambda_S l_B - \lambda_B l_S \leq \Lambda \gamma_+$) This bank can maximize either of Equations 11 or 12, under the constraint that $b_i = s_i$. We take first-order conditions and subsequently impose symmetry, such that each bank selects an identical b_i , s_i and H_i . The resulting solutions are $b_i = s_i = \frac{l_B + l_S}{\Lambda(N+1)}$ and $H_i = 0$. This solutions is optimal over any region where a bank wants to selects neither $b_i > s_i$, nor $s_i > b_i$.

B. Proof of Proposition 2

Proof follows from Proof of Proposition 1, making replacements of Γ_+ and Γ_- for γ_+ and γ_- , respectively.

C. Proof of Proposition 3

In the baseline case, banks trade on agency ($b_i = s_i$) when $-\Lambda\gamma_- \leq \lambda_S l_B - \lambda_B l_S \leq \Lambda\gamma_+$.

Under liquidity regulation, the same is true when $-\Lambda\Gamma_- \leq \lambda_S l_B - \lambda_B l_S \leq \Lambda\Gamma_+$. Since $\gamma_+ < \Gamma_+$ and $\gamma_- < \Gamma_-$, the region in which banks agency trade becomes larger.

D. Proof of Proposition 4

When a bank takes a long position, its net position in the baseline is given by $\frac{l_B - \gamma_+}{\lambda_B(N+1)} - \frac{l_S + \gamma_+}{\lambda_S(N+1)}$, while its net position under liquidity regulation is given by $\frac{l_B - \Gamma_+}{\lambda_B(N+1)} - \frac{l_S + \Gamma_+}{\lambda_S(N+1)}$. Since $\gamma_+ < \Gamma_+$, the net position declines.

E. Proof of Proposition 5

Gross quantity in the baseline case is given by:

$$N * (b_i + s_i) = \begin{cases} \frac{\lambda_S l_B + \lambda_B l_S + (\lambda_S - \lambda_B)\gamma_-}{\lambda_B \lambda_S (N+1)} & \text{Short} \\ \frac{2(l_B + l_S)}{\Lambda(N+1)} & \text{Agency} \\ \frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S)\gamma_+}{\lambda_B \lambda_S (N+1)} & \text{Long} \end{cases} \quad (27)$$

Gross quantity in the liquidity case is given by:

$$N * (b_i + s_i) = \begin{cases} \frac{\lambda_S l_B + \lambda_B l_S + (\lambda_S - \lambda_B)\Gamma_-}{\lambda_B \lambda_S (N+1)} & \text{Constrained Short} \\ \frac{2(l_B + l_S)}{\Lambda(N+1)} & \text{Agency} \\ \frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S)\Gamma_+}{\lambda_B \lambda_S (N+1)} & \text{Constrained Long} \end{cases} \quad (28)$$

Consider banks who would have taken a long position in the baseline ($\lambda_S l_B - \lambda_B l_S > \Lambda \gamma_+$). If they also take a long position with liquidity regulation ($\lambda_S l_B - \lambda_B l_S > \Lambda \Gamma_+$), it can be shown that:

$$\frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S) \Gamma_+}{\lambda_B \lambda_S (N + 1)} > \frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S) \gamma_+}{\lambda_B \lambda_S (N + 1)} \quad (29)$$

if $\lambda_B > \lambda_S$, and gross quantity increases. Otherwise, if $\lambda_S > \lambda_B$, gross quantity decreases.

Alternatively, if they do not take a long position with liquidity regulation ($\Lambda \gamma_+ \leq \lambda_S l_B - \lambda_B l_S < \Lambda \Gamma_+$) it can be shown that:

$$\frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S) \Gamma_+}{\lambda_B \lambda_S (N + 1)} > \frac{2(l_B + l_S)}{\Lambda(N + 1)} \quad (30)$$

if $\lambda_B > \lambda_S$, and gross quantity increases. Otherwise, if $\lambda_S > \lambda_B$, gross quantity decreases.

The proof is symmetric for banks who would have taken a short position in the baseline. However, gross quantity increases with regulation if $\lambda_S > \lambda_B$, or decreases if $\lambda_B > \lambda_S$.

Gross quantity is unchanged if banks take an agency position in the baseline.

F. Proof of Proposition 6

By computing prices and spreads with liquidity regulation, it can be shown that they remain at $\frac{l_B + l_S}{N + 1}$ regardless of whether banks take long, short or agency positions or the cost of any regulation.

G. Proof of Proposition 7

By inspection, for any given $\lambda_S l_B - \lambda_B l_S$, each part of Equation 17 is weakly lower after substituting Γ_- for γ_- and Γ_+ for γ_+ , since $\Gamma_- > \gamma_-$ and $\Gamma_+ > \gamma_+$. For example, in the short region,

$$\frac{\beta \Phi}{\lambda_S (N + 1)} l_S - \frac{\beta \Phi}{\lambda_B (N + 1)} l_B - \frac{\Lambda \beta \Phi}{\lambda_B \lambda_S (N + 1)} \Gamma_- < \frac{\beta \Phi}{\lambda_S (N + 1)} l_S - \frac{\beta \Phi}{\lambda_B (N + 1)} l_B - \frac{\Lambda \beta \Phi}{\lambda_B \lambda_S (N + 1)} \gamma_-.$$

H. *Proof of Proposition 8*

Ex-ante welfare is measured by the total expected surplus of sellers ($\frac{1}{2} * N * (P_B - v + l_B) * b_i$), buyers ($\frac{1}{2} * N * (v + l_B - P_S) * s_i$), and banks ($N * (v - P_B) * b_i + N * (P_S - v) * s_i$).

Consider banks who take long positions in the baseline ($\lambda_S l_B - \lambda_B l_S > \Lambda \gamma_+$). Total surplus is given by:

$$\frac{N}{N+1} \left(\frac{l_B(l_B - \gamma_+)}{\lambda_B} + \frac{l_S(l_S + \gamma_+)}{\lambda_S} \right) - \frac{N^2}{2(N+1)^2} \left(\frac{(l_B - \gamma_+)^2}{\lambda_B} + \frac{(l_S + \gamma_+)^2}{\lambda_S} \right) \quad (31)$$

If banks also take a long position with liquidity regulation ($\lambda_S l_B - \lambda_B l_S > \Lambda \Gamma_+$), total surplus is given by,

$$\frac{N}{N+1} \left(\frac{l_B(l_B - \Gamma_+)}{\lambda_B} + \frac{l_S(l_S + \Gamma_+)}{\lambda_S} \right) - \frac{N^2}{2(N+1)^2} \left(\frac{(l_B - \Gamma_+)^2}{\lambda_B} + \frac{(l_S + \Gamma_+)^2}{\lambda_S} \right) \quad (32)$$

Algebraic manipulation shows that when $\lambda_S l_B - \lambda_B l_S > \Lambda \Gamma_+$, Equation 32 is less than Equation 31, and thus, surplus is lower.

Alternatively, consider banks who would have taken a long position in the baseline, but who take an agency position with liquidity regulation ($\Lambda \gamma_+ \leq \lambda_S l_B - \lambda_B l_S < \Lambda \Gamma_+$). With liquidity regulation, total surplus in this case is given by:

$$\frac{N(l_B + l_S)^2}{\Lambda(N+1)} - \frac{N^2}{2\Lambda(N+1)^2} (l_B + l_S)^2 \quad (33)$$

Algebraic manipulation shows that when $\Lambda \gamma_+ = \lambda_S l_B - \lambda_B l_S$, then Equation 33 is equal to Equation 31. Alternatively, if $\Lambda \gamma_+ < \lambda_S l_B - \lambda_B l_S$, then Equation 33 is less than Equation 31. Thus, in either case, if banks would have taken a long position ($b_i > s_i$) in the baseline, then welfare falls with liquidity regulation.

The proof for banks who take short positions follows similar to the one above. Welfare is unchanged if banks take agency positions in the baseline.

I. Proof of Proposition 9

The leverage constraint limits banks' abilities to take a long position, such that $b_i - s_i \leq \Psi$. The bank's problem is unchanged from the baseline if it wishes to take a short or agency position.

Consider a bank which wishes to take a long position. The firm selects b_i , s_i and H_i to maximize Equation 11, such that $b_i - s_i \leq \Psi$. We take first-order conditions and subsequently impose symmetry, such that each bank selects an identical b_i , s_i and H_i .

If the constraint is slack, such that $b_i - s_i < \Psi$, the solutions are unchanged from the baseline $b_i = \frac{l_B - \gamma_+}{\lambda_B(N+1)}$, $s_i = \frac{l_S + \gamma_+}{\lambda_S(N+1)}$ and $H_i = 0$.

If the constraint binds, such that $b_i - s_i = \Psi$, the solutions to the bank's optimization problem are $b_i = \frac{l_B + l_S}{\Lambda(N+1)} + \frac{\lambda_S \Psi}{\Lambda}$, $s_i = \frac{l_B + l_S}{\Lambda(N+1)} - \frac{\lambda_B \Psi}{\Lambda}$ and $H_i = 0$. Algebraic manipulation shows that the constraint binds when $\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi \leq \lambda_S l_B - \lambda_B l_S$.

J. Proof of Proposition 10

A bank which becomes leverage constrained has equilibrium supplies $b_i = \frac{l_B + l_S}{\Lambda(N+1)} + \frac{\lambda_S \Psi}{\Lambda}$ and $s_i = \frac{l_B + l_S}{\Lambda(N+1)} - \frac{\lambda_B \Psi}{\Lambda}$ such that $b_i - s_i = \Psi$. In the baseline case, this same bank has liquidity supplies $b_i = \frac{l_B - \gamma_+}{\lambda_B(N+1)}$ and $s_i = \frac{l_S + \gamma_+}{\lambda_S(N+1)}$.

Algebraic manipulation can show that when a bank is bound by the leverage constraint $(\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi \leq \lambda_S l_B - \lambda_B l_S)$,

$$\Psi \leq \frac{l_B - \gamma_+}{\lambda_B(N+1)} - \frac{l_S + \gamma_+}{\lambda_S(N+1)} \quad (34)$$

When a bank is not leverage constrained, the equilibrium is identical to the baseline and liquidity supply is unchanged.

K. Proof of Proposition 11

Consider banks who take a long position in the baseline. Gross quantity traded is given by $\frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S) \gamma_+}{\lambda_B \lambda_S (N+1)}$. When banks are leverage constrained, gross quantity traded is given by $\frac{2(l_B + l_S)}{\Lambda(N+1)} + \frac{(\lambda_S - \lambda_B) \Psi}{\Lambda}$.

Algebraic manipulation can show that, when banks are leverage constrained ($\Lambda \gamma_+ + \lambda_B \lambda_S (N+1) \Psi \leq \lambda_S l_B - \lambda_B l_S$),

$$\frac{\lambda_S l_B + \lambda_B l_S + (\lambda_B - \lambda_S) \gamma_+}{\lambda_B \lambda_S (N+1)} > \frac{2(l_B + l_S)}{\Lambda(N+1)} + \frac{(\lambda_S - \lambda_B) \Psi}{\Lambda} \quad (35)$$

when $\lambda_S > \lambda_B$, and gross quantity traded decreases with regulation. Otherwise, if $\lambda_B > \lambda_S$, gross quantity traded increases with regulation.

If banks are not leverage constrained, the equilibrium is identical to the baseline and gross quantity traded is unchanged.

L. Proof of Proposition 12

By computing prices and spreads with leverage regulation, it can be shown that they remain at $\frac{l_B + l_S}{N+1}$ regardless of whether banks take long, short or agency positions or the cost of any regulation.

M. Proof of Proposition 13

When banks are leverage constrained, their returns are given by $r_A = \beta \Psi \Phi$. If banks are not subject to leverage regulation, their returns for long positions are given by $r_A = \beta \frac{\lambda_S l_B - \lambda_B l_S - \Lambda \gamma_+}{\lambda_B \lambda_S (N+1)} \Phi$. Algebraic manipulation shows that when banks are leverage constrained ($\Lambda \gamma_+ + \lambda_B \lambda_S (N+1) \Psi \leq \lambda_S l_B - \lambda_B l_S$),

$$\beta \Psi \Phi < \beta \frac{\lambda_S l_B - \lambda_B l_S - \Lambda \gamma_+}{\lambda_B \lambda_S (N+1)} \Phi \quad (36)$$

When banks are not leverage-constrained, the equilibrium is identical to the baseline and returns are unchanged.

N. Proof of Proposition 14

Consider banks who are leveraged constrained ($\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi \leq \lambda_S l_B - \lambda_B l_S$).

Total surplus is given by,

$$\frac{N}{\Lambda(N+1)} \left((l_B + l_S)^2 + (N+1)(\lambda_S l_B - \lambda_B l_S)\Psi \right) \quad (37)$$

$$- \frac{N^2}{2\Lambda^2} \left(\lambda_B \left(\frac{l_B + l_S}{(N+1)} + \lambda_S \Psi \right)^2 + \lambda_S \left(\frac{l_B + l_S}{(N+1)} - \lambda_B \Psi \right)^2 \right) \quad (38)$$

Welfare in the baseline, minus welfare in the leverage constrained case is given by subtracting Equation 37 from Equation 31. After some algebraic manipulation, this can be shown to be equal to,

$$\begin{aligned} \frac{N(\lambda_S l_B - \lambda_B l_S)}{2(N+1)^2 \Lambda \lambda_B \lambda_S} & \left((N+2)(\lambda_S l_B - \lambda_B l_S) - 2\Lambda\gamma_+ - 2(N+1)^2 \lambda_B \lambda_S \Psi \right. \\ & \left. + \frac{N(N+1)^2 \lambda_B^2 \lambda_S^2 \Psi^2 - \Lambda^2 N \gamma_+^2}{\lambda_S l_B - \lambda_B l_S} \right) \end{aligned} \quad (39)$$

First, consider banks that are at the exact point where they become welfare constrained ($\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi = \lambda_S l_B - \lambda_B l_S$). By substituting $\lambda_S l_B - \lambda_B l_S$ for $\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi$ in Equation 39, the welfare change can be shown to be zero.

Alternatively, if $\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi < \lambda_S l_B - \lambda_B l_S$, the bank is limited by its leverage constraint. By substituting $\lambda_S l_B - \lambda_B l_S$ for $\Lambda\gamma_+ + \lambda_B\lambda_S(N+1)\Psi + \epsilon$ in Equation 39, Equation 39 can be shown to be positive for any $\epsilon > 0$. Thus, if banks are leverage constrained, welfare is higher in the baseline.

If banks are not leverage constrained, the equilibrium is identical to the baseline and welfare is unchanged.