Intermediary market power and capital constraints*

Jason Allen[+] and Milena Wittwer[§]

January 25, 2023

Abstract

We study if and how intermediary capitalization affects asset prices in a framework that allows for intermediary market power. We show that weaker capital requirements lead to higher prices (lower yields) but greater markups due to market power. We test these predictions and calibrate the model with data on Canadian Treasury auctions, where we can link asset demand to balance sheet information of individual intermediaries. Our findings imply that weaker capital constraints lead to higher auction revenues and thus savings for the government at an implicit cost of larger yield distortion.

Keywords: Financial intermediaries, market power, price impact, asset demand, asset pricing, government bonds, Basel III, capital requirements, leverage ratios

JEL: G12, G18, G20, D40, D44, L10

---

*Previous versions of the paper were circulated under the title "Intermediary asset pricing: Capital constraints and market power". The views presented are those of the authors and not necessarily those of the Bank of Canada. Any errors are our own. We thank the Triton team at the Bank of Canada, Rui Albuquerque, Eric Budish, David Cimon, Darrell Duffie, Mark Egan, Liran Einav, Matthew Gentzkow, Daniel Smith, Or Shachar, Monika Piazzesi, Marzena Rostek, Xavier Vives, Ji Hee Yoon, in addition to all participants of seminars at Boston College, University of Maryland, Stanford, Boston Conference on Markets and Competition, European Winter Meeting of the Econometric Society, NBER Market Design, NYU IO day and the Annual Meeting of Women in Microstructure. Correspondence to: [+]Jason Allen (Bank of Canada): jallen@bankofcanada.ca and [§]Milena Wittwer - Boston College, Email: wittwer@bc.edu.

1 Introduction

What moves asset prices is one of the oldest questions in finance. The intermediary asset pricing literature suggests that the prices of many assets depend not only on the preferences of households, but also on the equity capitalization of financial intermediaries, called dealers (e.g., Brunnermeier and Pedersen (2009); He and Krishnamurthy (2012, 2013); Brunnermeier and Sannikov (2014)). One reason is that dealers face capital constraints. At the same time, dealers enjoy market power in various markets, including Treasury, repo and foreign exchange markets (e.g., Allen and Wittwer (2022); Huber (2022); Wallen (2022)). This stands in contrast to the traditional view that financial markets are highly competitive, motivating the assumption of perfect competition in models of intermediary asset pricing.

We zoom in on the intermediary sector to study how dealer capitalization affects asset demand, and therefore prices, and quantify the effect in a framework that allows for dealer market power (as in Kyle (1989)). We first introduce a model in which strategic dealers face regulatory capital constraints to highlight how their asset-demand, and ultimately prices, are impacted by their degree of capitalization and market power. Then, we provide evidence in favor of our model and calibrate it with data on Canadian Treasury auctions by leveraging regulatory changes during the COVID-19 pandemic.

In the model, presented in Section 2, dealers compete to buy multiple units of an asset that pays out an uncertain return in the future. They are risk averse and subject to a capital constraint. The market may clear via different auction formats, which represent primary auctions or exchanges. In the benchmark auction, each dealer submits a demand curve that specifies how much it is willing to pay for different units of the asset. The market clears at the price at which aggregate dealer demand meets supply, and each dealer wins the amount it asked for at that price (uniform price auction).

We show that there is an equilibrium in which each dealer's demand depends on whether and how tightly the capital constraint binds. This is measured by the shadow cost of the capital constraint. The main prediction of the model is that demand (which map prices into quantities) becomes flatter when capital constraints are relaxed as it becomes cheaper to buy larger amounts of the asset. Unless supply adjusts, the market price increases.[1] At the same time, dealers distort the market price further away from the price that would arise if they were price-takers. This is a source of inefficiency, for instance, because distorted prices in the primary market may distort trading and the security allocation in the secondary market.

We use data on Canadian Treasury auctions to test and calibrate the model leveraging two attractive features. First, dealers submit entire demand curves. We can therefore observe whether demand is flat or steep. Alternatively, we would need to aggregate individual demands from secondary market trades—which involves observing individual trades, and pooling data points over time and across market participants. Second, we can link the dealers' demand curves to balance sheet information, which is crucial for establishing a link between dealer demand and capitalization.

Our data, presented in Section 3, combines bidding information on all regular Canadian Treasury auctions between January 2015 and February 2022 with balance sheet information of the eight largest dealers at the company holding level (following He et al. (2017)). We observe all winning and losing bids. To make bonds with different coupons and maturities more comparable, bids are expressed in the yield-to-maturity (the annualized interest rate that equates the price with the present discount value of the bond). Therefore, we conduct our empirical analysis with yields instead of prices. In addition, we see the Basel III Leverage Ratio (LR) of each dealer-bank, which is the Canadian equivalent to the Supplementary Leverage Ratio (SLR) in the U.S. It is reported quarterly, measures a bank's Tier 1 capital relative to its total leverage exposure, and must be above an institutional-specific regulatory

[1]The prediction that the asset price increases when the shadow cost of the capital constraint decreases (because the constraint becomes weaker) is in line with He and Krishnamurthy (2012, 2013) and Brunnermeier and Sannikov (2014)'s prediction that a positive shock to the net worth, i.e., equity capital, of a dealer increases its risk-bearing capacity, which leads to higher asset prices. In our model, risk aversion is constant.

threshold, which we also observe.

With the data, we gather evidence in favor of the model's predictions in Section 4. For this, we leverage the feature that domestic government bonds were temporarily exempted from the LR to buffer against negative effects of the COVID-19 pandemic. Through the lens of our model, this means that the capital constraint was temporarily lifted. An event study confirms that the demand of dealers who were more strongly affected by the policy change became flatter when the exemption period started, and steeper when it ended, relative to the other dealers, as predicted by our model. In theory, this leads to a decrease in market yields and an increase in markups; empirically, these effects are hard to identify.

We leverage our model to quantify by how much the market yield and markups change when the constraint is relaxed or tightened in Section 5. To pin down the banks' degree of risk aversion and the shadow cost of the capital constraint, we first estimate how much bidders are willing to pay in each auction using estimation techniques from the auctions literature (introduced by Guerre et al. (2000); Hortaçsu and McAdams (2010); Kastl (2011)). With this, we are able to identify the degree of dealer risk-aversion and their shadow cost of capital by leveraging how dealer willingness to pay varies around the policy change, conditional on market observable characteristics (such as market volatility).

We find that banks are close to risk-neutral. This is reassuring given that these are global banks who can diversity away most types of risk. At the same time, banks face sizable costs due to the regulatory capital constraint. The shadow cost parameters imply that yields decrease and markups increase by 0.23%–0.49% if the shadow cost of the constraint decreases by 1%. In comparison, estimates from regressing auction yields on LR—which we suspect to be downward biased due to endogeneity—imply that a 1% increase in the LR (which decreases the shadow cost) decreases the yield by roughly 1.1%–1.6%. This result provides some validation to the model, given that the elasticities are not dissimilar and the bias goes in the expected direction, even though we have not used this information to estimate our model.

Our findings have valuable implications for government bond markets, which were under severe distress in March 2020. Globally, banking regulators took measures to facilitate central bank interventions to support financial intermediation, such as (temporarily) relaxing capital requirements.[2] Our counterfactual analyses highlights new welfare benefits and costs: the relaxation of the capital constraint leads to a reduction in bond yields but an increase in markups. This translates into higher auction revenues and thus savings for the government at an implicit cost of larger yield distortion.

Beyond implications for government bond markets, our evidence contributes to an ongoing debate in the asset pricing literature on whether, and if so how, capitalization of financial intermediaries affects asset prices (e.g., Adrian et al. (2014); He et al. (2017); Gospodinov and Robotti (2021)). We complement existing studies that rely on market-level data and proxy variables for intermediary costs by zooming in on one market in which we can observe the relationship between intermediary capitalization, demand and the market price. Since our model is not specific to the Canadian Treasury market, we conjecture that our findings generalize to many financial markets in which financial institutions intermediate trade.

Finally, our analysis contributes to the theoretic literature on intermediary asset pricing. To draw the connection, we extend our baseline model in Section 6 to study the extent to which intermediary market power affects whether intermediary frictions (such as moral hazard and capital constraints) matter for asset prices. Our findings highlight that it is important to take intermediary market power into account, and that price effects depend on the degree of market power. Hopefully, this motivates future research that can analyze the implications of intermediary market power in a macroeconomic model of intermediary asset pricing, and empirical research to assess the degree of competition in different financial markets.

[2]Capital requirements are meant to strengthen the risk management of banks and avoid the build up of systemic risks. Our analysis does not incorporate how these risks change when relaxing constraints.

Related literature. Our main contribution is to explain and quantify how dealer capitalization affects asset demand and prices when dealers have market power.

The topic fits into an ample intermediary asset pricing literature that examines the impact of dealer capitalization (or leverage) on asset price behavior due to constraints on debt (e.g., Brunnermeier and Pedersen (2009); Pedersen and Gârleanu (2011); Adrian and Shin (2014); Moreira and Savov (2017); Elenev et al. (2021)), or constraints on equity (e.g., He and Krishnamurthy (2013, 2012); Brunnermeier and Sannikov (2014)). Given our focus on banks, we follow He et al. (2017) and rely on equity constraints.

The key difference relative to these models is that we zoom in on the intermediary financial sector (for most of the paper) and allow dealers to impact prices as a result of market power in the tradition of Kyle (1989). The market clears via a multi-unit auction (as in Vayanos (1999); Vives (2011); Rostek and Weretka (2012); Malamud and Rostek (2017); Wittwer (2021)). Our innovation is to introduce a capital constraint and analyze its effect on market outcomes and market power.[3]

Our empirical analysis adds to a growing literature on the relation between intermediary costs or constraints and asset prices (e.g., Adrian and Shin (2010); Ang et al. (2011); Adrian et al. (2014); He et al. (2017, 2022); Du et al. (2018); Check et al. (2019); Gospodinov and Robotti (2021); Haddad and Muir (2021); Baron and Muir (2022); Fontaine et al. (2022)). Most existing studies use market-level data, such as cross-sectional returns of different asset classes, and rely on proxy variables to capture intermediary costs, such as the VIX or aggregate capital holdings.

We zoom in on one market in which we can link dealer demand with balance sheet

[3]Our extended model, presented in Section 6, is more similar to the intermediary asset pricing literature, which abstracts from market power with few recent exceptions (e.g. Corbae and D'Erasmo (2021); Jamilov (2021); Villa (2022); Wang et al. (2022)). In contrast to these papers, we analyze market power between intermediaries rather than market power of intermediaries vis-a-vis firms or consumers. Therefore, our insights, especially those on the linkage between intermediary market power and capital constraints, are fundamentally different from those found in this literature.

information to establish a direct relationship between dealer capitalization and asset demand, and identify a mechanism through which capital affects asset prices. Furthermore, we calibrate our model to quantify elasticities and conduct counterfactuals. For this, we rely on estimation techniques of a literature on multi-unit auctions, developed by Guerre et al. (2000), Hortaçsu and McAdams (2010) and Kastl (2011) and extended by Hortaçsu and Kastl (2012), and Allen et al. (2020, 2023). Our findings support Albuquerque et al. (2022) who suggest that the price elasticity of aggregate demand in Portuguese Treasury auctions may proxy for dealers' risk-bearing capacity.

Convention. Throughout the paper, we refer to markups as the difference between the price at which the market would clear if it was perfectly competitive and the price at which it clears under imperfect competition; or equivalently, as the difference between the yield at which the market clears under imperfect competition versus perfect competition. In a uniform price auction, markups increase in price impact—a common object of interest.

2 Model

Our goal is study how the market price and markups change when capital constraints are relaxed under different market clearing mechanisms. In our benchmark, we model market-clearing via a uniform price auction, which may be one-sided, meaning that bidders buy but not sell, or double-sided, so that bidders buy and sell. In practice, many primary markets, for instance in the U.S., clear via uniform price auctions, while trading on an exchange can be approximated via a double-sided uniform price auction (e.g., Kyle (1989)).

There are $N > 2$ dealers who compete in a uniform price auction to buy units of an asset of total supply $A > 0$; payment is in cash (numeraire). Dealer i holds $z_i \in \mathbb{R}$ of inventory of the asset, equity capital $E_i > 0$ as well as other assets on the balance sheet, which we normalize to 0 w.l.o.g. One unit of the asset pays a return of $R \sim N(\mu, \sigma)$ in the future. In our empirical application, where the asset is a government bond, R represents the return from selling the bond in the secondary market, which is unknown at the time of the auction.

Dealers may be uncertain about the supply, expected return of the asset, the inventory positions of other dealers or capital positions. In the simplest version of the model, presented here, dealers are uncertain about the supply A, thus face aggregate uncertainty, but have no private information about the asset's expected return μ, inventory z_i, or capital positions E_i. In Appendix A we generalize our results to the case of private information. In our empirical application, supply is random because dealers don't know the issuance size when they compete. In other settings, the supply might be random due to noise traders.

Each dealer submits a continuous and strictly decreasing demand schedule: $a_i(\cdot) : \mathbb{R} \to \mathbb{R}$, which specifies how many units of the asset, $a_i(p)$, the dealer seeks to buy at price, p. For a given price, the dealer chooses demand to maximize the utility it expects to earn from the wealth, $\omega_i(a_i, p)$, that would be generated if the dealer won amount a_i at price p, subject to a capital constraint:

$$\max_{a_i} \mathbb{E}_R \Big[1 - \exp\Big(- \rho_i \omega_i(a_i, p)\Big)\Big] \text{ subject to } \kappa_i \leq \frac{E_i}{p(a_i + z_i)}. \tag{1}$$

Parameter $\rho_i > 0$ measures the dealer's degree of risk-aversion. Future wealth, $\omega_i(a_i, p)$, is equal to the asset payoff, R, net of the price paid, p:

$$\omega_i(a_i, p) = (a_i + z_i)R - pa_i. \tag{2}$$

The capital constraint is motivated by the Basel III requirement according to which a bank must hold sufficient equity capital, E_i, relative to its total balance sheet exposure. In our case, the total exposure is the nominal amount of the asset the bank holds after the auction, $p(a_i + z_i)$.

Once all dealers have submitted their demand curves, the auction clears at the price, p^*, at which aggregate demand meets total supply, $\sum_i a_i(p^*) = A$, and each bank pays the market clearing price for the amount won, $a_i^* = a_i(p^*)$. After auction clearance, all transactions take place and the asset pays out its return.

We focus on Bayesian Nash equilibrium with linear demand curves (hereafter, equilibrium), which is common in the related literature (e.g., Kyle (1985, 1989), Vayanos (1999); Vives (2011); Malamud and Rostek (2017); Wittwer (2021)). Crucially, this does not imply that dealers can only submit linear demand curves. Instead, it is optimal for dealers to submit linear demand curves when all others do so.

Proposition 1. *(i) In equilibrium dealer i submits demand curve*

$$a_i(p) = ((1 + \lambda_i\kappa_i)\Lambda_i + \sigma\rho_i)^{-1}(\mu - (\sigma\rho_i + \Lambda_i\lambda_i\kappa_i)z_i - (1 + \lambda_i\kappa_i)p), \tag{3}$$

with $\Lambda_i = \beta_i\alpha_i\sigma$ with $\beta_i = \frac{2}{\alpha_i b - 2 + \sqrt{(\alpha_i b)^2 + 4}}$ and $\alpha_i = \frac{\rho_i}{1+\lambda_i k}$, where $b \in \mathbb{R}^+$ is the unique positive solution to $1/2 = \sum_i (\alpha_i b + 2 + \sqrt{(\alpha_i b)^2 + 4})^{-1}$, and the λ_i's are pinned down by the system of N equations: $\lambda_i(E_i - \kappa_i p^*(a_i(p^*) + z_i)) = 0$ $\forall i$ with $p^* : A = \sum_i a_i(p^*)$. This equilibrium exists for parameter constellations for which $\lambda_i \geq 0$ $\forall$ i.

(ii) When dealers are identical ($\rho_i = \rho, z_i = z, E_i = E, \kappa_i = \kappa$), the demand curve is

$$a_i(p) = \left(\frac{N-2}{N-1}\right)\frac{1}{\rho\sigma}\Big(\mu - \rho\sigma z - (1 + \lambda\kappa)p\Big), \tag{4}$$

with $\lambda = \frac{A}{NE}\left(\mu - \rho\sigma z - \sigma\rho\left(\frac{N-1}{N-2}\right)\frac{A}{N}\right) - \frac{1}{\kappa}$ when $E \neq \kappa p^*[a_i(p^*) + z]$ and $\lambda = 0$ otherwise. This equilibrium exists for parameter constellations for which $\lambda \geq 0$, for instance, when μ and κ are sufficiently high.

To derive an intuition for this equilibrium, consider a dealer who chooses how much to demand at price p. The dealer takes the behavior of the other dealers as given and maximizes its expected utility from winning the asset subject to market-clearing and the capital constraint. Optimal demand equalizes the expected marginal utility (LHS) with the marginal payment (RHS):

$$\mu - \rho_i\sigma(z_i + a_i) = (1 + \lambda_i\kappa_i)(p + \Lambda_i a_i) + \lambda_i\kappa_i\Lambda_i z_i. \tag{5}$$

The marginal utility is decreasing in the amount of the asset. For the first unit of the asset, the dealer earns the per-unit return. For the next units, the utility becomes smaller, depending on the variance of the asset's return and the dealer's degree of risk aversion.

The marginal payment has several components and depends on the regulatory shadow cost of the capital constraint (the Lagrange multiplier $\lambda_i \geq 0$), and the dealer's price impact, $\Lambda_i \geq 0$. The latter is known as Kyles' lambda and is 0 when the market is perfectly competitive so that dealers are price-takers. When the constraint is not binding ($\lambda_i = 0$) and dealers are price-takers ($\Lambda_i = 0$), the marginal payment is just the price, p, that the dealer has to pay for amount a_i. When the constraint binds ($\lambda_i > 0$) and dealers are price-takers ($\Lambda_i = 0$), the marginal payment is the price they have to pay plus a shadow cost that comes from the capital constraint, which is similar to an ad-valorem tax: $(1 + \lambda_i \kappa_i)p$. When dealers face a binding capital constraint ($\lambda_i > 0$) and have market power ($\Lambda_i \neq 0$), $\Lambda_i a_i$ measures by how much a dealer's choice impacts the effective price $(1 + \lambda_i \kappa_i)p$. Not only does this depend on their risk-aversion and the number of players in the market, it also depends on the shadow cost of capital. For instance, with identical dealers $\Lambda_i = \frac{1}{N-2}\frac{\rho\sigma}{1+\lambda\kappa}$ for all i. Finally, when $z_i \neq 0$, there is an extra term, $\lambda_i k \Lambda_i z_i$, which reflects the regulatory cost that comes from the fact that a dealer's existing inventory, z_i, is evaluated at the market price in the capital constraint.

The key prediction of the model is about what happens when the capital constraint is relaxed, for instance because the minimal capital thresholds decrease.

Corollary 1. *Let $p^*(0)$ denote the market price when dealers are price-takers and $p^*(\Lambda)$ when they have market power, and consider a relaxation of capital constraints which decreases all λ_i's, for instance, a decrease in the κ_i thresholds. The demand $a_i(\cdot)$ of each dealer i becomes flatter, and market price $p^*(\Lambda)$ increases. Further, the price impact Λ_i of each dealer i, and the markup $= p^*(0) - p^*(\Lambda)$ increase.*

Figure 1: Own and cross demand effect when capital constraints are relaxed

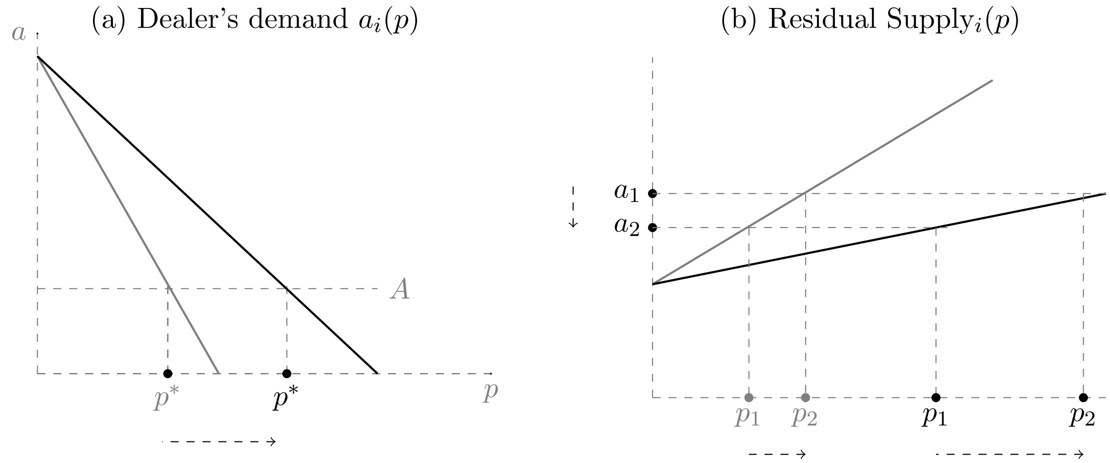


Figure 1 illustrates the change in the dealer's own demand and her residual supply curve when capital constraints are relaxed in (a) and (b), respectively, for the case in which no dealer carries inventory ($z_i = 0$ for all i). In gray we see the initial demand curve and residual supply curve. Both become flatter, as shown by the black line, when constraints are relaxed. In (a) we see how this increases the market clearing price, p^*, when supply is fixed. In (b) we see the increase in the price impact, which measures by how much the clearing price changes, $p_2 - p_1$, when the dealer marginally changes her demand from a_1 to a_2.

Figure 1 illustrates two effects from relaxing capital constraints. The first is an own-demand effect. Since the effective price, $(1 + \lambda_i\kappa_i)p$, decreases, it becomes cheaper for the dealer to buy larger amounts. The dealer's demand, $a_i(\cdot)$, flattens and the market price increases mechanically, unless supply adjusts.

The second effect comes from the change in demand of other dealers. This is because the relaxed capital constraint not only affects dealer i, but all other dealers as well. All other dealers submit flatter demand curves, which implies that the residual supply curve that dealer i faces when choosing its own demand schedule is flatter. A flatter residual supply curve, in turn, means that the dealer moves the market clearing price more strongly when changing her demand. The dealer's price impact and markup increase.

When dealers are identical, we can derive by how much the price and the markup change in response to a 1% change in the shadow cost of capital, i.e., $\lambda\kappa$.

Corollary 2. *Let dealers be identical. When the shadow cost of capital* ($\lambda\kappa$) *decreases by 1%, the market price and the markup increases each by* $\eta = \left|\frac{1}{1+\lambda\kappa} - 1\right|\%$.

Take away. Our model helps explain how capital constraints affect asset prices and markups. The main prediction is that both the market price and markups increase when capital constraints are relaxed, which highlights a trade-off for debt-managers in primary markets.

In Appendix A, we show that this prediction (captured in Corollaries 1 and 2) generalizes to discriminatory price auctions, in which winning bidders pay their bids, not the market clearing price. This auction format is used to sell government debt in many countries, including Canada.

3 Institutional setting and data

To test the predictions of our model (Corollary 1) and to quantify by how much demand, the price, or equivalently the yield, and markups change when capital constraints are relaxed (Corollary 2), we use data on Canadian Treasury auctions. They have the attractive feature that dealers submit entire demand curves, which we can link to balance sheet information of each dealer at the company holding level.

Market players. There are eight deposit-taking primary dealers in Canada who are federally regulated.[4] They dominate the Canadian Treasury market and intermediate the vast majority of the daily trade volume in government bonds. More broadly, these banks dominate the Canadian banking sector and hold over 90% of the sector's assets.

[4]In total there are eleven primary dealers. One of these dealers is provincially regulated and two are private securities dealers. They face different capital regulation than the eight dealers we study. We therefore do not observe any balance sheet information for these players. Technically, two of the eight banks have multiple dealers. For example, the Bank of Montreal has two dealers (Bank of Montreal and BMO Nesbitt Burns) who attend different Treasury auctions, and therefore do not compete or share information within an auction. We treat them as one dealer.

Primary dealers have a responsibility, as market-makers, to buy bonds from the government and trade them with investors, brokers, or one another to provide liquidity. They hold a substantial amount of bonds on their own balance sheets (see Appendix Figure A1). In exchange, primary dealers enjoy benefits, including privileged access to liquidity facilities and overnight repurchase operations at the central bank.

Market-making is a small part of the bank's total business, which includes accepting deposits, making loans, and wealth management. Therefore, the dealer has no control over most determinants on the bank's balance sheet. A dealer's assets are on average 9% of the bank's total assets (Allen and Usher (2020)). The bank's balance sheet, in turn, is what matters for the regulator since capital requirements must be met at the company holding level.

Treasury auctions. Governments issue bonds in the primary market via regularly held uniform price or discriminatory price auctions. In Canada, regular auctions are discriminatory price. They take place several days a week. Anyone may participate, but most of the supply is purchased by dealers. The largest eight dealers purchase on average 81% of the supply in order to sell (or lend) on the secondary market.[5]

Capital constraints. According to a survey among market participants, the Basel III LR represents the most relevant capital constraint when trading government bonds (CGFS (2016)). This regulatory requirement came into effect in September 2014 to reduce systematic risk—a benefit which we do not consider in this paper. We focus on the cost-side of the constraint, which was emphasized by Duffie (2018), He et al. (2022) and others.

Formally, the LR measures a bank's Tier 1 capital relative to its total leverage exposure, and must be at least 3%:

[5]This percentage represents how much of the amount that is issued to bidders other than the Bank of Canada, who bids non-competitively, is allocated to the eight largest dealers.

$$\mathrm{LR}_{iq} = \frac{\text{regulatory capital of bank } i \text{ in quarter } q}{\text{total leverage exposure of } i \text{ in } q}.$$

Tier 1 capital consists primarily of common stock and disclosed reserves (or retained earnings), but may also include non-redeemable non-cumulative preferred stock; the leverage exposure includes the total notional of all cash and repo transactions of all securities, including government bonds, regardless of which securities are used as collateral (for more details see OSFI (2023)).

In reality, banks refrain from getting close to the minimal Basel III threshold (see Figure 3b, explained below).[6] One reason for this is that each institution faces an additional supervisory LR threshold that reflects the underlying risk of the bank's operations. To determine this threshold, the regulator considers a range of factors, such as operating and management experience, strength of parent institution, earnings, diversification of assets, type of assets, and appetite for risk (Engert (2005)). These institution-specific thresholds are communicated to institutions on a bilateral basis, and are considered supervisory information. They are not permitted to be disclosed to the public (see OSFI (2023), Section IV). Another reason is that banks tend to hold sufficient conservation buffer for Tier 1 capital so as to avoid punishment in the form of restricted distributions (including dividends and share buybacks, discretionary payments and bonus payments to staff).

Data. The unique feature of our data is that we can link how a dealer bids in the Treasury auction to balance sheet information about the dealer's bank.

We obtain bidding data of all regular Treasury auctions between January 1, 2015 and February 1, 2022 from the Bank of Canada. We see who bids (identified by a legal entity identifier) and all winning and losing bids. For consistency, we restrict attention to bids of the eight dealers who are deposit-taking throughout most of the paper.

[6] Barth et al. (2005), Berger et al. (2008) and Brewer et al. (2008) document that bank capital is substantially above the regulatory minimum in countries other than Canada.

We collect balance sheet information for these eight dealers at the bank level. Specifically, we obtain the LRs and supervisory LR thresholds of each bank. Both are reported quarterly, at the end of January (first quarter), April (second quarter), July (third quarter) and October (fourth quarter of the reporting year) from January 2015 until January 2022 from a data source, called LR.[7] In addition, we obtain the daily aggregated long and short positions in government bonds of the six largest dealer-banks from the Collateral and Pledging Report (H4). Finally, we collect information on who holds government bonds—banks versus other investor types—from the National Accounts (Statistics Canada).

In addition, we obtain yields of all trades with Canadian government bonds in the secondary market from November 2015 until December 2020 to better isolate the yield effect of dealer capitalization. These data are collected by the Industry Regulatory Organization of Canada in the Debt Securities Transaction Reporting System (MTRS2.0) and are made available for research with a time lag.

Finally, we collect the Implied Volatility Index for Canadian Treasuries from the Bank of Canada. The index measures the expected volatility of Treasury prices over the next 30 days, similar to the Merrill Lynch Option Volatility Estimate (MOVE) for U.S. Treasuries or CBOE Volatility Index (VIX) for stocks. It is based on option prices on interest rate futures (Chang and Feunou (2014)).

Conventions and summary statistics. So far, we have used prices to express bond values, because it is more intuitive for us to think through the economics when demand schedules are downward sloping. From now on, we express bond values in yields-to-maturity, as is the data. This makes the value of bonds that have different maturities and coupon payments more comparable. Since yields increase when prices decrease, demand schedules with yields are upward sloping. An overview of the main variables is in Table 1.

[7] One of the banks, HSBC, has a different reporting schedule than the others. Its fiscal year ends in December, instead of October. In our empirical analysis this difference is absorbed when we include dealer fixed effects.

Table 1: Summary statistics

	Mean	Median	Std	Min	Max
Total amount issued (in bn C$)	3.97	3.20	2.33	1	20.6
Average bid yield (in %)	0.92	0.65	0.63	0.04	2.50
Days to maturity	1,094	352	2,292	99	12,399
Number of dealers per auction	8.00	8.00	0.04	7	8
Number of steps in demand curve	4.74	5	1.67	1	7
Maximal amount demanded (in mil C$)	741	600	569	0.48	8,240
Amount dealer won (in mil C$)	277	150	395	0	5,739
Quarterly LR (in %)	4.41	4.36	0.28	-	-
Implied Volatility Index (in %)	0.46	0.25	0.77	0.04	7.57

Table 1 shows the average, median, standard deviation, minimum and maximum of key variables in our sample. Our auction data goes from January 1, 2015 until February 1, 2022 and counts 917 auctions. In total there are 21 different securities. The min and max LR are empty because we cannot disclose this information.

4 Evidence in favor of our model

We now provide supporting evidence of our model. In Section 5 we calibrate the model to quantify by how much yields and markups change when capital constraints are relaxed.

Demand effect. To provide evidence that demand becomes flatter when capital constraints are relaxed, we leverage two features. First, when dealers failed to absorb the extraordinary supply of government bonds in March 2020, government bonds, central bank reserves, and sovereign-issued securities that qualify as High Quality Liquid Assets (HQLA) were temporarily exempted from the LR constraint—starting on April 9, 2020.[8] As a result,

[8]Exposures related to the US Government Payment Protection Program (PPP), which are minor in the case of Canadian banks, were also temporarily exempted. The announcement to start the exemption period is available at: `www.osfi-bsif.gc.ca/Eng/fi-if/in-ai/Pages/20200409-dti-let.aspx`; the one to end it is here: `www.osfi-bsif.gc.ca/Eng/fi-if/in-ai/Pages/lrfbunwd.aspx`, both accessed on 05/31/2022.

the LR spiked upward, moving away from the constraint (see Figure 3b). The exemption of government bonds and HQLA ended on December 31, 2021, while reserves continued to be excluded. Second, in the absence of the exemption some banks faced higher, i.e., stricter, capital thresholds than other banks. According to our model, the demand of these banks should have become flatter during the exemption period than the demand of other banks.

To show the predicted effect, we construct a measure of the slope of demand and analyze how slopes of banks with higher institution-specific capital thresholds changed relative to those of the other banks. Creating a slope measure is challenging because it must be dealer and auction-specific so as to account for differences in capitalization, capital requirements and market conditions. For instance, we cannot simply regress demand on prices or yields and use the regression coefficient as our slope measure. Not only are prices endogenous, but this approach would eliminate the variation across time and dealers that is crucial for estimation.

To construct a slope measure, assume for a moment that dealers can submit linear demand curves that map from prices to quantities, like in the model. Dealer i's demand on day t for security s would be: $a_{its}(p) = intercept_{its} - slope_{its} \times p$. If this was so, we could read the intercept and slope off the data (see Figure 2). In reality, there are two differences. The first is that dealers cannot submit linear demand curves but instead have to submit step functions with maximally 7 steps. A dealer i submits $K_{its} \leq 7$ quantity-price tuples: $\{a_{itsk}, p_{itsk}\}_{k=1}^{K_{its}}$. This implies that the measure we introduce,

$$slope_{its} = \frac{\max_k\{a_{itsk}\}}{\max_k\{p_{itsk}\}} \text{ for } K_{its} > 1, \tag{6}$$

is an approximation (see Figure 2). The second difference is that the value of the bond is expressed in yields rather than prices, which means that the submitted demand curves are upward sloping. To incorporate this, we convert the maximal price in equation (6) into the corresponding yield.

Figure 2: Example of a demand function with 6 steps

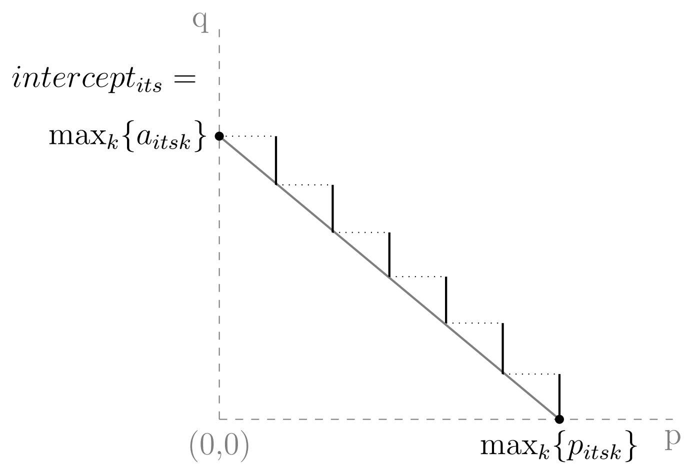


Figure 2 illustrates how we measure the slope of demand of dealer i on day t for security s. In black, we show an example of a step-function, $\{a_{itsk}, p_{itsk}\}_{k=1}^{6}$, that a bidder could submit at auction (with prices instead of yields). In gray, we show a continuous, linear demand function which connects the maximal amount that the bidder asks for, $\max_k\{a_{itsk}\}$, with the maximal price that it is willing to pay, $\max_k\{p_{itsk}\}$. We use the slope of this auxiliary function to approximate the "slope" of the step function, and convert prices into yields.

One concern with slope measure (6) is that it relies on two extreme points on the demand curve.[9] To ensure that our estimation results are not biased by extreme points, we perform robustness checks with alternative measures in Appendix D.

To conduct our event study, we regress the average slope in the demand function of dealer i in quarter q for security s on an indicator variable D_k equal to 1 k quarters before/after the first event quarter (2020q2) multiplied by the bank's supervisory LR threshold in that quarter.[10] We include quarter-security fixed effects, ζ_{qs}, to absorb common unobservables that affect dealer demand, such as the start of quantitative easing (QE) or COVID-related

[9]Ex ante, the dealer does not know where the market will clear, and submits bids at prices that may win with positive probability. However, ex post with finite data, it may be that we don't observe dealers ever winning at extreme points on their demand curves.

[10]We do not include dealer fixed effects because the thresholds are persistent over time with few changes. As part of our robustness analyses, we estimate an analogous regression with leverage ratios LR_{ik} replacing the $theshold_{ik}$ and with dealer fixed effects.

Figure 3: The effect of the exemption on Treasury positions and the LR

(a) Aggregated positions in Treasuries (b) Time series of LR for an average bank

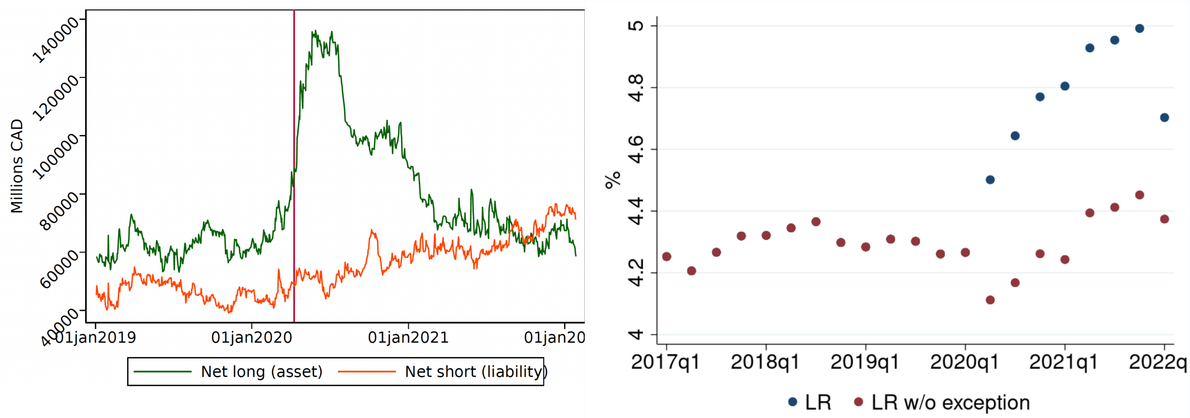


Figure 3a shows the aggregated amount of Canadian government bonds that the biggest six Canadian banks hold in long (in green) and short (in red) positions in millions of C\$ from January 2019 until February 2022. The vertical line is April 9, 2020, when government bonds were exempt from LR. Figure 3b shows the time series of the LR (in %) of an average bank. In blue, we show the actual LR. In red the counterfactual LR that the average bank would have had if government bonds, central bank reserves, sovereign-issued securities that qualify as HQLA, and exposures related to the PPP were not exempt. In 2022q1, the LR does not get back to its original level, partially because central bank reserves are still exempted.

demand and supply factors:

$$slope_{iqs} = \alpha + \sum_{k=-K}^{K} \gamma_k \times D_k \times theshold_{ik} + \zeta_{qs} + \epsilon_{iqs}. \tag{7}$$

The slope is measured in million C\$ per 1 bps. For, instance, a slope of 1 means that a dealer demands C\$ 1 million less when the yield increases by 1 bps. The threshold is also expressed in bps.

The parameter of interest, γ_0, tells us by how much the slope of a bank that faces a tighter constraint (higher supervisory LR threshold) changes relative to the slope of a bank that faces a less stringent constraint (lower supervisory LR threshold), when government bonds are exempted, and similarly for γ_7.

In line with our model, we find that the estimated coefficients are not statistically different

Figure 4: Change in the slope

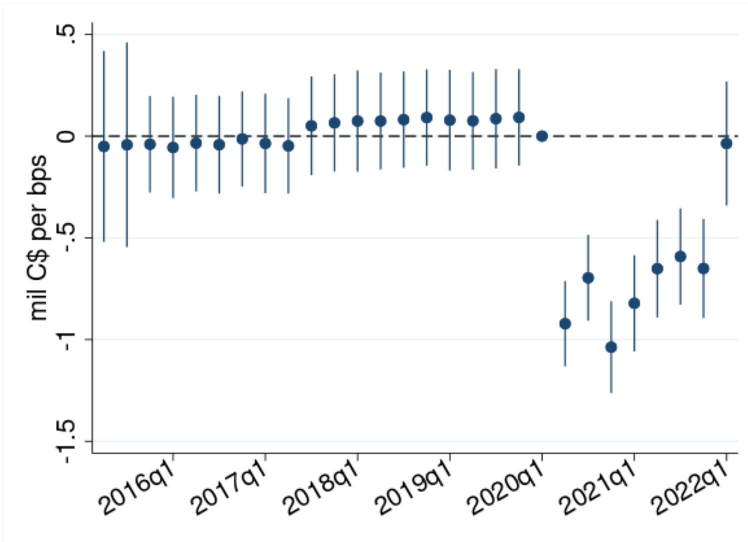


Figure 4 shows the γ_k estimates and 95% confidence intervals of the regression (7): $slope_{iqs} = \alpha + \sum_{k=-K}^{K} \gamma_k \times D_k \times threshold_{ik} + \zeta_{qs} + \epsilon_{iqs}$. All magnitudes are relative to the benchmark in 2020q1. The slopes are in million C$/bps, thresholds are in bps.

from 0 when government bonds are not exempt from the LR, but drop during the exemption period (see Figure 4). The size of the effect is large (C$1 million per bps). This suggests that, during the exemption period, a bank whose capital threshold is 1 bps above the threshold of another bank reduces its demand by C$ 1 million less than the other bank in response to a 1 bps yield increase. This change is sizable compared to the average amount a dealer wins at auction (C$ 277 million).

The estimated relationship would be causal, meaning that demands became flatter as a result of the exemption, under three conditions. First, bank-specific capital thresholds were not changed in response to, or set in expectation of, the temporary exemption. This is plausible given that COVID was unpredictable. Second, the Canadian regulator did not change the LR requirement in response to how dealers bid in the primary auctions—which is likely true given that the Canadian response to COVID was part of a global response (see Bank of International Settlement (2020)). Third, there were no systematic changes over time (other than the exemption) that affected the slope of different dealers differently.

Crucially, we do not require banks with tighter thresholds to be identical to those with weaker thresholds. For instance, it is okay for banks to differ in their underlying risk in operations. To confirm the model's prediction we only need that these unobserved factors do not change materially during the exemption period, and therefore explain the entire change in demand.

Two confounding factors come to mind. First, the Bank of Canada increased the maximal bidding limits from April 2020 until June 2021. This was to allow dealers to buy larger amounts of government bonds, even though pre-COVID maximal bidding limits were typically not binding (see Appendix Figure A2). Second, the Bank of Canada started buying government bonds in the secondary market via QE from April 2020 until October 2021. Both of these policy interventions were meant to help ensure that dealers continued to buy in the primary market.[11] Thus, if anything, these interventions likely increased the amount a dealer asked for at auction. By definition, such increases would have increased the slope measure (6). We find the opposite and therefore don't think that the change in bidding limits or QE drives the drop in 2020q2. Moreover, neither of the two confounding factors were present in 2022q1, when the slope jumps back up.

Yield effect. When demand becomes flatter the market clearing yield must decrease, all else equal. This is a mechanical effect in theory, which is hard to show in the data. The reason is that we only observe one market clearing yield per auction. We don't see the counterfactual scenario in which demand curves would have been flatter and the market would have cleared at a different yield, and can no longer rely on variation in thresholds across dealers. However, we can study how yields change when the LR increases—either

[11] This raises the concern that dealers were at the minimal bidding limits. These limits are only soft constraints since dealers are given 6 months to improve their performance if they fall below the minimum. They are bank-specific and the exact numbers are unknown to us. However, we know that for most banks the minimal limit is around 10% and do not see that banks are more likely to be at that limit during 2020–2021 (see Appendix Figure A2).

Table 2: Correlation between yield and LR

	(OLS)	(FE1)	(FE2)
LR	-0.360^{***}	-0.370^{**}	-0.245^{**}
	(0.036)	(0.041)	(0.052)
controls	–	–	✓
fixed effects	–	✓	✓
Observations	2,912	2,912	2,904
Adjusted R^2	0.032	0.679	0.789

Column (OLS) of Table 2 shows results of $yield_t = \alpha + \beta LR_{qi} + \epsilon_{ti}$. In (FE1) we add the dealer and year fixed effect; in (FE2) the control variables. Yield and LR are in %. Standard errors are in parentheses, clustered at the dealer level in (FE1) and (FE2). $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

because banks hold sufficient capital, which lowers their shadow cost of capital, or because capital requirements are relaxed—and control for factors that move yields but are unrelated to dealer capitalization.

For this we regress the yield in auction t on the quarterly LR of a dealer, LR_{qi}, control variables, $controls_t$, and dealer, ζ_i, as well as year fixed effects, ζ_y:

$$yield_t = \alpha + \beta LR_{qi} + controls_t + \zeta_y + \zeta_i + \epsilon_{ti}. \tag{8}$$

First, to take out unobservable factors that move yields (such as interest rate uncertainty and inflation risk) we control for the yield at which the bond with the closest maturity date is traded on that day in the secondary market; on average the maturity differs by a few days.[12] Second, to remove any effect that extra supply might have on the yield, we control for the supply of the auction.

In line with the model's prediction, we report in Table 2 a negative relationship between the auction yield and the LR. The average LR increase from 4.3% to 4.5% in 2020q1–2020q2

[12]One concern is that this bond is less liquid than the bonds issued due to an on-the-run effect. However, this effect is less common in Canada because many bonds are re-issued multiple times to avoid low liquidity of a particular security.

correlates with a yield drop of 5–7 bps, which is sizable given the low yields in the sample.

The estimates would be downward biased (and hence larger in absolute value than they should be) if the LR was positively correlated with the error term. This would be the case if there were unobservable factors that lead banks to decrease the LR and decrease the yield. Before April 9, 2020, one way for this to happen is if banks actively increased their LR in response to some negative capital shock by buying less government bonds in a way that decreases the total asset exposure but keeps capital constant. Then, since dealer demand is downward (upward) sloping in price (yield), the market yield would decrease.

Take away. Our empirical evidence provides support for our model hypothesis that dealer demand becomes flatter when capital constraints are relaxed. This decreases the market yield.

5 Quantification

To quantify by how much the yield and markups change when capital constraints are relaxed, we take our model to the data. The size of the change depends on the shadow cost of the capital constraint, and the dealer's risk aversion, in addition to observables, such as market volatility.

5.1 Identifying risk aversion and shadow costs

To estimate the shadow cost and the degree of risk aversion, we face three challenges that stem from the fact that Treasury auctions are more complicated in reality than what we can capture with a tractable theory. First, the auction is discriminatory price, not uniform price. Second, demand functions are step-functions. Third, dealers are not the only bidders in the auction. Customers (which include pension, mutual and ETF funds, insurance companies, sovereigns, and bank treasuries) can participate but must place their bids with a dealer who passes them to the auctioneer. This means that dealers might obtain information from observing their customer bids.

To overcome these challenges, we use insights and techniques from the empirical literature on auctions (Guerre et al. (2000): Hortaçsu and McAdams (2010); Kastl (2011); Hortaçsu and Kastl (2012); Allen et al. (2023)). The main idea is that we can back out how much bidders are truly willing to pay from the bids they submit under the assumption that bidders are rational and play an equilibrium of an auction game that incorporates the key institutional features of Canadian Treasury auctions. To do this, we do not have to solve for an equilibrium, but only characterize (necessary) equilibrium conditions. With the bidders' estimated willingness to pay—which depends on the shadow cost and risk aversion according to our theory—we can separately identify these parameters by leveraging the fact that Treasuries were temporarily exempt from the constraint.

Identifying assumptions. There are N_D potential dealers and N_C potential customers who participate in auction t. They play the auction game of Hortaçsu and Kastl (2012), which minics the auction process of Canadian Treasury auctions.

Before bidding, each dealer and customer draws a private signal ξ_{ti} about her true willingness to pay $v_{ti}(q)$ for amount q. The signal ξ_{ti} is drawn independently from all other bidders according to atomless distributional functions $F_t^D(\xi_{ti})$ and $F_t^C(\xi_{ti})$ for dealers and customers, respectively. This implies that dealers are ex-ante identical, but may be asymmetric relative to customers.[13]

A bidder's signal can be persistent over time, but it must be independent from all other bidders' signals conditional on any information that is known to all bidders at the time of the

[13]In theory, we could allow dealers with high capital thresholds to be different from those with low capital thresholds by creating a third bidder group. In practice, we quickly run into power issues without pooling information across auctions, which we refrain from to avoid estimation biases that come form unobservable auction characteristics. The presence of such characteristics implies correlations in bidder signals when going across auctions, which violates one of our identifying assumptions.

auction. This includes a price-range which is provided by the Bank of Canada for reference, and all public information that is available in the active when-issued and secondary market. Given the liquidity and structure of these markets—for example, the fact that many clients trade with the same dealer (Allen and Wittwer (2022))—it seems reasonable to assume that bidders are equally informed about the price at which they can sell the bond after the auction, unless they receive diverging private information, for instance, from observing client orders. To further validate the assumption of conditional independence, Hortaçsu and Kastl (2012) and Allen et al. (2023) conduct statistical tests and find supporting evidence.

Different to Hortaçsu and Kastl (2012), we assume that a dealer's willingness to pay, $v_{ti}(\cdot)$, has a specific functional form (9). To see why, recall equation (5). This equation characterizes how much a dealer is truly willing to pay, which is simply the price that the dealer would submit if she was a price-taker (i.e., $\Lambda_i = 0$). Adjusting the notation to highlight that parameters may change across auctions t, we rearrange equation (5) with $\Lambda_{ti} = 0$ to obtain the dealer's true willingness to pay for amount a,

$$v_{ti}(a) = \xi_{ti} - \left(\frac{\rho_t \sigma_t}{1 + \lambda_t \kappa_t} \right) a, \tag{9}$$

when dealers are (ex-ante) identical. According to our theory, $\xi_{ti} = (\mu_t - \rho_t \sigma_t z_{ti})(1 + \lambda_t \kappa_t)^{-1}$, although for estimation we do not have to impose any specific functional form.

To identify dealer risk aversion and the shadow cost of the capital constraint, we assume that the dealer's risk aversion does not change in auctions close to the policy change, $\rho_t = \rho_q$, and that the shadow cost is constant within a quarter, $\lambda_t \kappa_t = \lambda_q \kappa_q$. We include all auctions in the quarter before and after the two policy changes, motivated by the fact that capital requirements are only required to hold quarterly.

Estimation procedure Our estimation procedure has two steps.

First, we estimate the dealers' true willingness to pay, $\hat{v}_{tik} = \hat{v}_{ti}(q)$, for each submitted amount q, or, equivalently submitted step k, from the necessary equilibrium conditions (characterized in Kastl (2012), Proposition 1). These conditions depend on what dealers believe about where the auction will clear, that is, the distribution of the market clearing price. To estimate this distribution from the viewpoint of each bidder in each auction, we adopt the resampling method of Allen et al. (2023) who generalize Hortaçsu and Kastl (2012). The core idea is to draw from the submitted bids of an auction to simulate possible market outcomes.

Second, we leverage the assumption that a dealer's willingness to pay is given by equation (9), together with the fact that most dealers submit a step function with more than one step. We first estimate an auction-specific slope coefficient, β_t, by regressing the dealer's estimated value at step k on the amount she demanded at that step, a_{tik}, and a dealer-time-auction fixed effects, ξ_{ti},

$$\hat{v}_{tik} = \xi_{ti} + \beta_t a_{tik} + \epsilon_{tik}, \tag{10}$$

using values from dealers who submit more than one step in auction t.[14] Then, we normalize each slope coefficient by the Implied Volatility Index for Canadian Treasuries, shown in Figure 5a, which measures return volatility, σ_t, in the dealer's willingness to pay (9). The normalized coefficient, $\frac{\beta_t}{\sigma_t}$, is equal to the dealer's risk aversion, ρ_t, during the exemption period (in which the shadow costs are 0) and $\frac{\rho_t}{1+\lambda_t \kappa_t}$ otherwise.

Finally, under the assumption that risk aversion does not change around the policy change and that the shadow cost of capital is constant within a quarter, we can back out our parameters of interest by comparing slope coefficients around the two policy changes. Alternatively, we can estimate these parameters in one step with the following regression,

$$\hat{v}_{tik} = \zeta_{ti} + \rho \times exempt_t \sigma_t a_{tik} + \frac{\rho}{1 + \lambda\kappa} \times (1 - exempt_t)\sigma_t a_{tik} + \epsilon_{tik}, \tag{11}$$

using estimated values, $\hat{v}_{tik}$, and respective quantities a_{tik} of dealers who submit more than one step in an auction t during 2020q1 and 2020q2 when the exemption period started. For

[14] As robustness, we also use only functions with more than 2 or 3 steps. Findings are similar.

the end of the exemption period, we use data from 2021q4 and 2022q2. Indicator variable $exempt_t$ is 1 in 2020q1 and 2022q1, respectively, and 0 otherwise; σ_t is the daily Implied Volatility Index, and ζ_{ti} is a dealer-auction-time fixed effect.

5.2 Estimated parameters and elasticities

Risk aversion and shadow costs. As a first step, we visualize the distribution of $\frac{\beta_t}{\sigma_t}$ in Figure 5b. This is already informative. We see that ρ_t is small, with a median of around 0.001. This is reassuring since we would expect global banks to be able to diversity away most risks. In addition, we see that ρ_t varies across auctions. One reason for this is that different auctions offer government bonds of different maturities and risk aversion varies in the length to maturity. Longer bonds are more risky to buy than short T-bills. Finally, the median $\frac{\beta_t}{\sigma_t}$ outside of the exemption period is lower than the median during the exemption period, which suggests that the shadow cost is strictly positive in 2020q1 and 2022q2.

In line with these insights, we estimate low risk aversion and high shadow costs using regression (11)—see Table 3a. The risk aversion is +0.0001 in 2020q1–2020q1 and +0.0003 in 2021q4–2022q1, which means that a dealer's average willingness to pay decreases by only C\$2.77–C\$8.31 when going from winning nothing to winning the average amount (C\$277 million).[15] Dealers appear close to risk neutral, which would be the case if the willingness to pay was perfectly flat. The shadow cost of capital is very high—0.96 in 2020q1–2020q2, when dealers were asked to absorb large amounts of Treasuries onto their balance sheets, which tightened capital constraints. The cost decreases to 0.30 in 2021q1–2022q4 when the Treasury market had calmed down.

As a sanity check, we also estimate regression (11) with submitted, observed bids, instead

[15]Risk aversion for 2021q4–2022q1 is slightly higher than for 2020q1–2020q1, perhaps, because dealers were uncertain about when the exemption for settlement balances would be lifted. This would tighten their LR. In addition, dealers might have been concerned that the Bank of Canada would start unwinding their bond QE-purchases, which would increase their exposure and lower their LR given their obligation to buy these bonds.

Figure 5: Implied Volatility Index and normalized coefficients of regression (10)

(a) Implied Volatility Index: σ_t

(b) Normalized coefficients: $\frac{\hat{\beta}_t}{\sigma_t}$

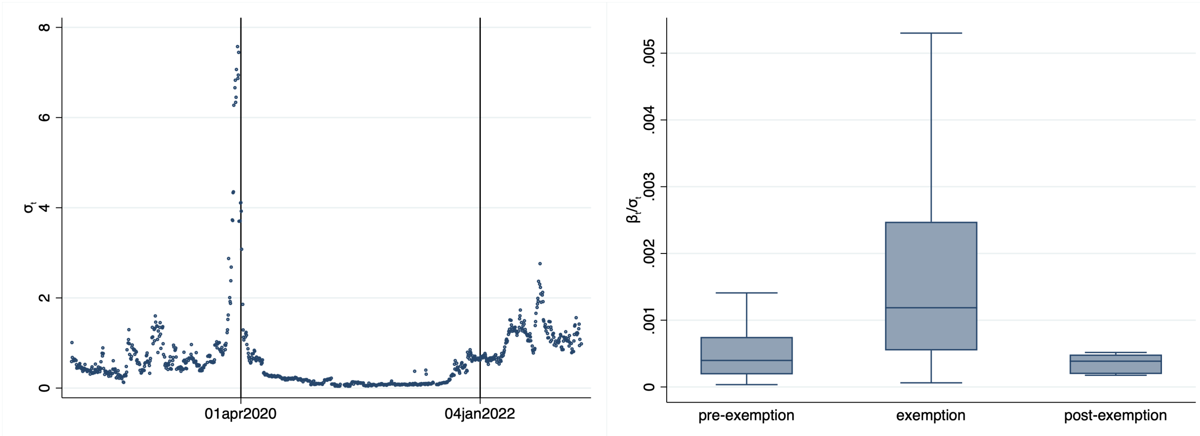


Figure 5a displays the time series of the daily Implied Volatility Index σ_t (in %) of the Canadian Treasury market from 2019 until March 2022. Figure 5b shows the distribution of the estimated slopes coefficient of the dealers' willingness to pay in auction t, $\hat{\beta}_t$ of regression (10) with dealer values in % and quantities in million C\$, normalized by σ_t for three time periods: before the exemption of Treasuries from the LR (2019q1–2020q1), during the exemption period (2020q1–2021q4) and after the exemption (2022q1).

of estimated values. In line with our expectations, parameter estimates in Table 3b are similar in magnitudes but not identical. In particular, estimates are slightly downward biased when using bids. The reason is that dealers shade early steps, i.e., small amounts, more strongly than higher steps, i.e., large amounts (see Figure 6). This implies that the dealer's true willingness to pay is steeper than the submitted bidding function.

The high shadow cost estimates suggest that it is not profitable for dealers to buy Treasuries in the primary market. To see this, consider a dealer in 2020q1–2020q2 who holds zero inventory going into an auction. To buy amount a of bonds, a dealer must pay weakly more than p^*a in a discriminatory price auction plus $0.96p^*a$ due to the capital constraint. In return, the dealer earns Ra from selling the bonds in the secondary market. For this transaction to be profitable it must be that the dealer can sell the bond at a price that is at least double the auction price in the secondary market. This never happens in the data.

Table 3: Estimated risk aversion and shadow cost parameters from regression (11)

(a) With values as independent variable

	2020q1-2020q2		2021q4-2022q1	
ρ	$+1.52/10^4$***	$(0.033/10^4)$	$+3.96/10^4$***	$(0.155/10^4)$
$\lambda\kappa$	+0.965***	(0.187)	+0.302**	(0.115)
N	23,074		12,894	

(b) With bids as independent variable

	2020q1-2020q2		2021q4-2022q1	
ρ	$+0.68/10^4$***	$(0.010/10^4)$	$+2.04/10^4$***	$(0.045/10^4)$
$\lambda\kappa$	+0.842***	(0.141)	+0.169*	(0.068)
N	23,074		12,894	

Tables 3a shows the estimates of regression (11) $\hat{v}_{tik} = \zeta_{ti} + \rho \times exempt_t \sigma_t q_{tik} + \frac{\rho}{1+\lambda\kappa} \times (1 - exempt_t)\sigma_t q_{tik} + \epsilon_{tik}$ with values expressed in yields in %, and quantities in million C\$. In Table 3b we replace the estimated values by the observed bids b_{tik}. Standard errors are in parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

This raises the question of why any dealer would want to participate in Treasury auctions. It suggests that it must be valuable for banks to have the primary dealer status, otherwise we should observe more dealers exiting the market. Given that primary dealers are carefully chosen by the government and heavily regulated, the status signals trustworthiness and stability. This helps to attract bank clients who seek to invest outside of the government bond market. Primary dealers also have access to central bank liquidity facilities and can use their position in the government bond market to cross-sell to investors other investment products, such as underwriting or trading corporate debt. When the benefit of having the primary dealer status no longer out-weights the cost, dealers exit the market. This has happened several times over the past decades (Allen et al. (2023)).

Figure 6: Distribution of bid shading per step k

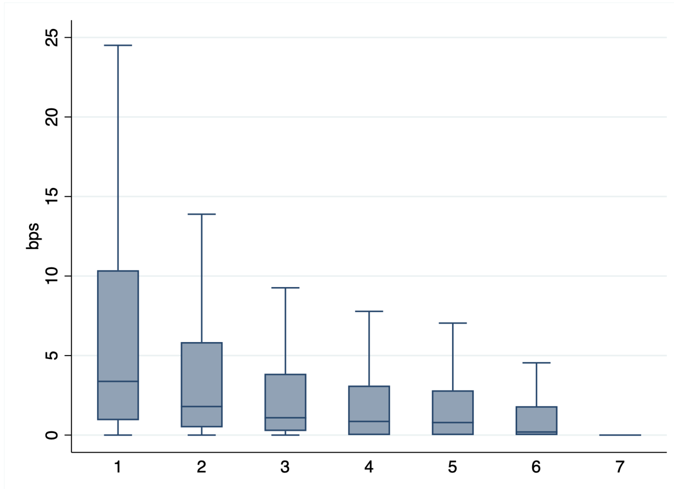


Figure 6 shows box plots of how much dealers shade their bids at each of the seven steps. Formally, it is the difference between the submitted yield bid and the estimated value, both in bps. The distribution for each step is taken over dealers and auctions. Shading factors are small in absolute terms, and comparable to those in the literature (e.g., Chapman et al. (2007); Kang and Puller (2008); Kastl (2011); Hortaçsu et al. (2018); Allen et al. (2020, 2023)).

Elasticities. With the estimated risk aversion parameters and shadow costs, we can approximate by how much yields and markups change when capital constraints are relaxed.

For this, we leverage Corollary 2, which generalizes to discriminatory price auctions with ex-ante identical dealers that have private information and submit linear demand schedules (see Appendices A and C). In practice, demand schedules are step-functions. However, they are approximately linear, in that the R^2 of regression (10) with bids instead of values is high (see Appendix Table A1). Therefore, we can rely on Corollary 2 to approximate by how much the market price (yield) increases (decreases) and the markup increases when the shadow cost of capital decreases by 1%, namely by $\eta = \left|\frac{1}{1+\lambda\kappa} - 1\right|\%$.

The shadow cost point estimates presented in Table 3 imply an η of 0.49% and 0.23% in 2020q1–2020q2 and 2021q4–2022q1, respectively. This means that the market yield decreases and the markup increases by 0.23%–0.49% if the shadow cost decreases by 1%.

In comparison, the regression evidence, presented in Table 2, says that a 1 pp, or, equivalently, a $1/4.41 \approx 23\%$ increase of the average LR—moving banks further away from the capital thresholds and thus decreasing the shadow costs—decreases the yield by roughly 25–36 bps. This implies that a 1% increase in the average LR decreases the average yield (which is 92 bps) by roughly $25/23$–$36/23 = 1$–1.5 bps, or equivalently 1.1%-1.6%. These numbers are in the same ballpark, but slightly larger, than the calibrated yield effect—in line with our conjecture that the regression coefficients might be biased downward because of endogeneity. This provides a validation for our model since we haven't used the regression information to estimate it.

Counterfactual. With the model we can quantify by how much the auction yield and markups would have changed had the regulator not changed the LR requirement.

For illustration, consider 1Y-bond auctions around both policy changes. The first auction after the LR was relaxed (on April 14, 2020) cleared at a yield of 51 bps. The average amount by which a dealer shaded her bid at market clearing, which approximates the markup due to market power, was roughly 2 bps. Had the LR not been relaxed—which implies a 100% increase in the shadow cost of the capital constraint—the auction would have cleared at a yield of $(1{+}0.49)51 \approx 76$ bps with a markup of $(1{-}0.49)2 \approx 1$ bps. The first auction after the LR was tightened (on January 5, 2022) cleared at 89 bps, with average shading of roughly 1 bps at market clearing. Had the LR not been tightened, the auction would have cleared at $(1{-}0.23)89 \approx 68$ bps with markup of $(1{+}0.23)1 \approx 1.2$ bps

This highlights a new trade-off that regulators should take into account when deciding whether to relax or tighten capital constraints. Relaxing capital constraints decreases yields but increases markups—in addition to the way the LR affects trading in the secondary market and concerns about systematic risk. Whether the effect is sizable in absolute terms depends on the interest rate level. If interest rates are low—as is the case in our sample—the absolute effect is small. For instance, a 1 bps yield increase per auction translates into an annual revenue loss of approximately C$ 69.3 million in 2021. However, with raising interests rates and growing balance sheets of banks, the trade-off may become first order in the future.

6 Implications for intermediary asset pricing

The main focus of this paper is on analyzing the effect of changing capital constraints on the price of an asset and markups that arise due to dealer market power. To draw a closer connection to the intermediary asset pricing literature and inspire future research, we extend our formal analysis in Appendix B to study how intermediary market power affects whether intermediary frictions (such as moral hazard or capital constraints) matter for asset prices. Here we only briefly mention the main take aways from this exercise and refer to the appendix for details.

Our first finding highlights that intermediary market power matters for intermediary asset pricing (see Corollary 4). We show that it is not possible to eliminate both moral hazard frictions and frictions that arise due to imperfect competition by hiring a manager who competes for the asset and is paid a fraction of the return that the asset will generate. In this sense, intermediary financing frictions always affect the asset price when the asset market is imperfectly competitive.

Our second finding highlights that capital constraints affect the asset price differently depending on how competitive the asset market is (see Corollary 5). We show that when intermediary market power increases because fewer intermediaries compete for the asset, frictions that arise from capital constraints distort the price more or less strongly, depending on asset-market competition. To be more concrete, consider a market in which many intermediaries compete for the asset. If the number of intermediaries decreases, the asset price moves further away from the price that would arise without capital constraints. The reason is that each intermediary wins more of the asset when fewer of them compete. This increases total exposure, and tightens the capital constraint. The opposite is true in a market with few intermediaries. Now, even though each intermediary wins more, the less competitive auction clears at a sufficiently low price. The price effect dominates the quantity effect and relaxes the constraint.

Taken together, these findings underline that it matters to take imperfect competition into account when analyzing how intermediary frictions affect asset prices, and motivates future research to assess the different degrees of competition across asset markets.

7 Conclusion

This paper studies if and how the capitalization of dealers affects asset demand and, with that, prices, when dealers have market power. We introduce a model to show that weaker capital requirements lead dealers to demand more of the asset at higher prices but also higher markups. We test the model's prediction and calibrate the model with data on Canadian Treasury auctions, where we can link asset demand to balance sheet information of individual intermediaries. Our findings highlight that weaker capital requirements reduce the funding cost of debt but increase market power.

References

Adrian, T., Etula, E., and Muir, T. (2014). Financial intermediaries and the cross-section of asset returns. *The Journal of Finance*, 69(6):2557–2596.

Adrian, T. and Shin, H. S. (2010). Liquidity and leverage. *The Journal of Financial Intermediation*, 19:418–437.

Adrian, T. and Shin, H. S. (2014). Procyclical leverage and value-at-risk. *The Review of Financial Studies*, 27(2):373–403.

Albuquerque, R., Cardoso-Costa, J. M., and Faias, J. A. (2022). Price elasticity of demand and risk-bearing capacity in sovereign bond auctions. Working paper.

Allen, J., Hortaçsu, A., Richert, E., and Wittwer, M. (2023). Dealer exit and hedge fund entry in the Canadian Treasury market. Mimeo.

Allen, J., Kastl, J., and Wittwer, M. (2020). Interconnected multi-unit auctions: An empirical analysis. Working paper.

Allen, J. and Usher, A. (2020). Investment dealer collateral and leverage procyclicality. *Empirical Economics*, 58(2):489–505.

Allen, J. and Wittwer, M. (2022). Centralizing over-the-counter markets? Bank of Canada working paper No. 2021-39.

Ang, A., Gorovyy, S., and Inwegen, G. B. (2011). Hedge fund leverage. *Journal of Financial Economics*, 102(1):102–126.

Ausubel, L. M., Cramton, P., Pycia, M., Rostek, M., and Weretka, M. (2014). Demand reduction and inefficiency in multi-unit auctions. *The Review of Economic Studies*, 81:1366–1400.

Bank of International Settlement (2020). Governors and heads of supervision announce deferral of Basel III implementation to increase operational capacity of banks and supervisors to respond to Covid-19. Press release.

Baron, M. and Muir, T. (2022). Intermediaries and asset prices: International evidence since 1870. *The Review of Financial Studies*, 35(5):2144–2189.

Barth, J., Caprio, G., and Levine, R. (2005). *Rethinking Bank Regulation: Till Angels Govern*. Cambridge University Press.

Berger, A. N., DeYoung, R., Flannery, M. J., Lee, D., and Öztekin, Ö. (2008). How do large banking organizations manage their capital ratios? *Journal of Financial Services Research*, 34:123–149.

Brewer, E., Kaufman, G. G., and Wall, L. D. (2008). Bank capital ratios across countries: Why do they vary? *Journal of Financial Services Research*, 34:177–201.

Brunnermeier, M. K. and Pedersen, L. H. (2009). Market liquidity and funding liquidity. *The Review of Financial Studies*, 22(6):2201–2238.

Brunnermeier, M. K. and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, 104(2):379–421.

CGFS (2016). CGFS papers no 55: Fixed income market liquidity. *Bank of International Settlements*.

Chang, B. and Feunou, B. (2014). Measuring uncertainty in monetary policy using implied volatility and realized volatility. *Bank of Canada Review*, pages 32–41.

Chapman, J. T. E., Mc-Adams, D., and Paarsch, H. J. (2007). Bounding revenue comparisons across multi-unit auction formats under ε-best response. *American Economic Review*, 97(2):455–458.

Check, H., Joslin, S., and Ni, S. (2019). Demand for crash insurance, intermediary constraints, and risk premia in financial markets. *The Review of Financial Studies*, 32(228-265).

Corbae, D. and D'Erasmo, P. (2021). Capital buffers in a quantitative model of banking industry dynamics. *Econometrica*, 86(6):2975–3023.

Du, W., Tepper, A., and Verdelhan, A. (2018). Deviations from covered interest rate parity. *The Journal of Finance*, 73(3):915–957.

Duffie, D. (2018). Financial regulatory reform after the crisis: An assessment. *Management Science*, 64:4835–4857.

Elenev, V., Landvoigt, T., and Nieuwerburgh, S. V. (2021). A macroeconomic model with financially constrained producers and intermediaries. *Econometrica*, 89(3):1361–1418.

Engert, W. (2005). On the evolution of the financial safety net. *Financial System Review*, pages 67–73.

Fontaine, J., Garcia, R., and Gungor, S. (2022). Intermediary leverage shocks and funding conditions. Working paper.

Gospodinov, N. and Robotti, C. (2021). Common pricing across asset classes: Empirical evidence revisited. *Journal of Financial Economics*, 140:292–324.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574.

Haddad, V. and Muir, T. (2021). Do intermediaries matter for aggregate asset prices? *The Journal of Finance*, 76(6):2719–2761.

He, Z., Kelly, B., and Manela, A. (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35.

He, Z. and Krishnamurthy, A. (2012). A model of capital and crises. *The Review of Economic Studies*, 79(2):735–777.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *American Economic Review*, 103(2):732–70.

He, Z. and Krishnamurthy, A. (2018). Intermediary asset pricing and the financial crisis. *Annual Review of Financial Economics*, 10:173–97.

He, Z., Nagel, S., and Song, Z. (2022). Treasury inconvenience yields during the COVID-19 crisis. *Journal of Financial Economics*, 143(1):57–79.

Holstrom, B. and Tirole, J. (1997). Financial intermediation, lonable funds, and the real sector. *The Quarterly Journal of Economics*, 112(3):663–691.

Hortaçsu, A. and Kastl, J. (2012). Valuing dealers' informational advantage: A study of Canadian Treasury auctions. *Econometrica*, 80(6):2511–2542.

Hortaçsu, A., Kastl, J., and Zhang, A. (2018). Bid shading and bidder surplus in the U.S. Treasury auction. *American Economic Review*, 108(1):147–69.

Hortaçsu, A. and McAdams, D. (2010). Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish Treasury auction market. *Journal of Political Economy*, 118(5):833–865.

Huber, A. W. (2022). Market power in wholesale funding: A structural perspective from the triparty repo market. Working paper.

Jamilov, R. (2021). A macroeconomic model with heterogeneous banks. Working paper.

Kang, B.-S. and Puller, S. L. (2008). The effect of auction format on efficiency and revenue in divisible goods auctions: A test using Korean treasury auctions. *The Journal of Industrial Economics*, 56(2):290–332.

Kastl, J. (2011). Discrete bids and empirical inference in divisible good auctions. *The Review of Economic Studies*, 78:974–1014.

Kastl, J. (2012). On the properties of equilibria in private value divisible good auctions with constrained bidding. *Journal of Mathematical Economics*, 48:339–352.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.

Kyle, A. S. (1989). Informed speculation with imperfect competition. *The Review of Economic Studies*, 56(3):317–355.

Malamud, S. and Rostek, M. (2017). Decentralized exchange. *American Economic Review*, 107(11):3320–3362.

Moreira, A. and Savov, A. (2017). The macroeconomics of shadow banking. *The Journal of Finance*, 72(6):2381–2432.

OSFI (2023). Leverage requirements guideline. `http://www.osfi-bsif.gc.ca/Eng/fi-if/rg-ro/gdn-ort/gl-ld/Pages/LR22.aspx`, accessed on 01/13/2023.

Pedersen, L. H. and Gârleanu, N. (2011). Margin-based asset pricing and deviations from the law of one price. *The Review of Financial Studies*, 24(6):1980–2022.

Rostek, M. and Weretka, M. (2012). Price inferences in small markets. *Econometrica*, 80(2):687–711.

Vayanos, D. (1999). Strategic trading and welfare in a dynamic market. *The Review of Economic Studies*, 66:219–254.

Villa, A. T. (2022). Credit misallocation and macro dynamics with oligopolistic financial intermediaries. Working paper.

Vives, X. (2011). Strategic supply function competition with private information. *Econometrica*, 76(6):1919–1966.

Wallen, J. (2022). Markups to financial intermediation in foreign exchange markets. Working paper.

Wang, Y., Whited, T. M., Whu, Y., and Xiao, K. (2022). Bank market power and monetary policy transmission: Evidence from a structural estimation. *The Journal of Finance*, 77(4):2093–2141.

Wittwer, M. (2018). *Pay-as-bid auctions in theory and practice.* PhD thesis, European University Institute.

Wittwer, M. (2021). Connecting disconnected financial markets? *American Economic Journal: Microeconomics*, 13(1):252–282.

Appendix

Overview. Appendix A generalizes our benchmark model to incorporate private information in uniform price and discriminatory price auctions. Appendix B generalizes the model to draw implications for the intermediary asset pricing literature. Proofs are in Appendix C. Robustness analyses are in Appendix D.

A Model extension: Private information

In our benchmark model, potentially asymmetric dealers face aggregate uncertainty in a uniform price auction. In the empirical model, dealers are ex-ante identical but have private information. Here we show that the model's prediction generalize to such a setting. In reality, dealers could be uncertain about the asset's mean return, their inventory, or equity positions. We focus on private information about inventory positions. The other cases are analogous.

Uniform price auction. There are two differences compared to the benchmark model. First, dealers are ex-ante identical ($E_i = E, \kappa_i = \kappa, \rho_i = \rho \ \forall i$). Each draws an inventory position z_i independently from all other dealers from some distribution. A dealer observes her own position privately prior to the auction, in addition to the aggregate inventory position, $\sum_i z_i$. Second, dealers balance out their total exposure post-auction in a market outside of the model, for example in an inter-dealer market.[16] This assumption reflects the fact that by the end of quarter banks want to balance their exposure. Further, it makes the model tractable because it ensures that dealers can predict the Lagrange multipliers, λ_i, of their competitors.[17]

[16] For simplicity, balancing is done at an exogenous price, normalized to zero.

[17] With unknown λ_i's there is no linear equilibrium, which substantially complicates the analysis. As comparison, the related literature only considers linear equilibria.

We model the balancing process in reduced form by assuming that for each z_i and market price p there is an $x_i(p)$ such that $a_i(p)+z_i+x_i(p)=\eta(p)$. Here $a_i(\cdot)$ is the dealer's demand schedule, and $\eta(\cdot)$ is a mapping from the price to the total exposure a dealer carries when the capital constraint must be satisfied. In reality this happens at the end of each quarter. To achieve tractability within the linear-demand environment, we assume that $\eta(p)$ is linear, i.e., $\eta(p)=\alpha-\beta p$ for some $\alpha\in\mathbb{R}^+,\beta\in\mathbb{R}$. We focus on the case in which exposure decreases in price, $\beta>0$. An equilibrium for the other case can be derived analogously.

The sequence of events is as follows: First, each dealer observes her inventory position, z_i, and the aggregate inventory, $\sum_i z_i$. Then, each dealer submits her demand schedule, $a_i(\cdot)$. Finally, the auction clears at price $p^*:\sum_i a_i(p^*)=A$, and dealers balance out their total exposure, so that their exposure equals $\eta(p^*)$. The asset pays its return and all transactions are made.

Proposition 2. *In equilibrium dealer i submits*

$$a_i(p)=(\Lambda+\sigma\rho)^{-1}\Big(\mu-\alpha\kappa\Lambda\lambda-\sigma\rho z_i-(1-\beta\kappa\Lambda\lambda)p\Big) \tag{12}$$

with $\Lambda=\frac{(N-2)\pm\sqrt{(N-2)^2-4\beta\kappa\lambda(N-1)\rho\sigma}}{2\beta\kappa\lambda(N-1)}$ *and λ such that* $\lambda(E-\kappa p^*\eta(p^*))=0$ *where* $p^*:\sum_i a_i(p^*)=A$. *This equilibrium exists as long as $\lambda\geq 0$, $\Lambda>0$, and $(1-\beta\kappa\Lambda\lambda)>0$.*

Like in the benchmark model, each dealer chooses how much to demand at each price p so as to equate marginal utility and marginal payment: $\mu-\rho\sigma(z_i+a_i)=p+\Lambda_i a_i+\Lambda_i\lambda_i\kappa\eta(p)$. The marginal payment is slightly different from before because the dealer now anticipates that she holds $\eta(p)$ of total exposure on her balance sheet at the end of the game. The constraint now binds when $E=\kappa p\eta(p)$ rather than when $E=\kappa p(a_i+z_i)$.

Discriminatory price auction. Here we adjust our benchmark model to fit discriminatory price auctions, in which bidders pay the prices they offered to pay for all units they win rather than the market clearing price. We make two changes to the benchmark setting. First, we rely on the fact that dealer i's true willingness to pay for amount a is

$v_i(a) = \frac{1}{1+\lambda\kappa}[\mu - \sigma\rho(z_i + a)]$ given objective function (1) and shadow cost of capital $\lambda \geq 0$.[18] Here, z_i is the dealer's private information, which is drawn iid across i from a distribution with support $[\underline{z}, \overline{z}]$. Supply A is drawn from some distribution with support $[\underline{A}, \bar{A}]$.

Second, we assume that the amount that a bidder wins in the auction follows a Generalized Pareto Distribution.[19] Formally, winning quantities are drawn from a distribution with CDF $F_i(a) = 1 - (\frac{\nu_i + \xi a}{\nu_i})^{-\frac{1}{\xi}}$ with $\xi \in (-\infty, -1], \nu_i = -\xi(\frac{N(1-\xi)-1}{N(1-\xi)})(\overline{z} - z_i) - \xi(\frac{\bar{A}}{N})$. Proving existence and characterizing an equilibrium for discriminatory price auctions for other distributions of winning quantities (when bidders have private information and demand multiple units) is a challenging and open question in the literature, that is beyond the scope of this paper.

Proposition 3. *There exists an equilibrium in which dealer i submits*

$$a_i(p) = \left(\frac{N(1-\xi)-1}{N-1}\right)\frac{1}{\sigma\rho}\left(\mu_i - (1+\lambda\kappa)p\right) \tag{13}$$

with $\mu_i = \mu + \rho\sigma(\frac{1}{1-\xi}(\xi\overline{z} - z_i) + \frac{\xi\bar{A}}{N(1-\xi)-1})$ for quantities that she wins with positive probability.

The intuition for this equilibrium is similar to the intuition for the equilibrium in a uniform price auction. The main difference is that bidders form different expectations over how likely it is to win each unit.[20]

Importantly, all equilibria of Propositions 1-3 have a similar functional form. The only difference between discriminatory price and uniform price auction is that the price impact is not clearly defined for the former. We therefore omit the price impact in the following corollary (which generalizes Corollary 1).

[18] We explain why this is true in Section 5.1

[19] This assumption implies functional form restrictions on the model's exogenous primitives that cannot be easily summarized.

[20] See Ausubel et al. (2014) who describe the trade-off in a framework without private information.

Corollary 3 (Analogue to Corollary 1)**.** *When the capital constraint is relaxed so that* λ *decreases, demand becomes flatter, and the market price and markup increase in a uniform and discriminatory price auction.*

B Intermediary asset pricing implications

Here we study whether the degree of competition between intermediaries affects the way intermediary frictions—specifically, moral hazard and capital constraints—affect asset prices. Technically, we rely on a simple version of the He and Krishnamurthy (2012, 2013) models, presented in He and Krishnamurthy (2018), that builds on Holstrom and Tirole (1997). Our contribution is to introduce imperfect competition in the asset market by relying on insights from the literature on auctions and market microstructure.

Model with moral hazard. The economy runs for three periods, $t = 0, 1, 2$.[21] There is one risky asset of aggregate supply A that pays out a return R per unit and a numeraire (cash). The supply and return of the asset are unknown to all agents in all periods but the last one; $R \sim N(\mu, \sigma)$ and A is drawn from some distribution.

There are $2 < N < \infty$ intermediaries, that is banks, indexed by I. Each bank serves a unit mass of households (H) who never consider switching banks. You may therefore think of fixed households-bank pairs. In addition, each bank has a trading desk i who is responsible for trading the risky asset.

Households cannot directly invest in the asset market, but must invest via a bank. For this, the households and their bank contract with trading desk i (of the bank that serves the households) who invests on the households' behalf in the risky asset.

Banks and households have CARA preferences, that is, holding wealth ω_j generates the following utility for an agent of type $j \in \{H, I\}$:

$$u_j(\omega_j) = 1 - \exp\left(-\rho_j \omega_j\right), \tag{14}$$

[21] Alternatively we could merge $t = 0$ and $t = 1$ into a single period.

with risk aversion $\rho_j > 0$. A trading desk and its bank share the same utility function. Wealth comes from buying and holding the asset. For instance, if agent j gets a_j at price p, the wealth is $\omega_j = a_j(R - p)$.

The sequence of events is as follows: In period 0 each households-bank pair chooses what fraction ϕ_i of the total wealth (that will be generated from investing in the asset) will be paid to trading desk i in period 2. The contract is chosen to maximize their joint expected utility obtained at the end of the game.[22] In period 1, all N trading desks compete in a uniform price auction to buy a_i of the risky asset. As in the baseline model presented in Section 2, bidders (here trading desks) face uncertainty about the asset's supply when placing their bids, but have no private information. Different to before, each trading desk may decide to shirk or exert effort; $s_i \in \{0, 1\}$, where s_i=1 is shirking. When a trading desk chooses not to exert effort, the wealth of its bank falls by Δ, but the trading desk gains a private benefit of b. In period 2, the asset pays its return and supply realizes. All transactions take place.

Proposition 4. *Define $m = \frac{\Delta}{b} - 1 \geq 0$. There exists an equilibrium in which $\phi_i = \phi = \frac{1}{1+m}$, and the clearing price is*

$$p^* = \mu - \left(\frac{\sigma \rho_I}{1+m}\right)\left(\frac{N-1}{N-2}\right)\frac{A}{N}. \tag{15}$$

Trading desk i buys amount $\frac{A}{N}$, its bank obtains $\phi\frac{A}{N}$, and each mass of households $(1-\phi)\frac{A}{N}$.

In this equilibrium, the market clearing price has the familiar functional form of a uniform price auction with N bidders (here trading desks). $m = \frac{\Delta}{b} - 1$ is the maximum amount of dollars that households can invest (per dollar that the trading desk purchases) so that the

[22] You may think of a market designer who chooses all contracts to maximize expected welfare of the economy subject to incentive constraints. If asset markets were perfectly competitive, we could let the pair choose an affine contract parametrized by (K_i, ϕ_i) where ϕ_i is the linear share of the return generated by the fund (conditional on supply A) that is paid to the trading desk, and K_i is a management fee that is paid to the trading desk and is independent of the fund's return. In the case of CARA preferences, the lack of a wealth effect implies that K_i plays no role in asset demand and equilibrium prices.

trading desk exerts effort in the auction. If the moral hazard friction is small, which happens when the benefit b from shirking is small, the trading desk can be incentivized to exert effort with little skin in the game, that is, with a small ϕ_i. The more beneficial it becomes to shirk, the higher ϕ_i must be.

When the asset market is perfectly competitive as in He and Krishnamurthy (2018), there are two cases depending on how attractive it is for the trading desk to shirk. In the first case, shirking is attractive so that the constraint that incentivizes trading desk i to exert effort, $\phi_i \Delta \geq b$, binds, or equivalently $\phi_i = \frac{1}{1+m}$. As a result, the intermediation frictions affect the asset price. In the second case, the incentive constraint doesn't bind, and the first best solution can be obtained through choosing the optimal contract ϕ_i.

When the asset market is imperfectly competitive, intermediation frictions always affect the asset price. Intuitively, this is because one instrument (per households-bank pair), ϕ_i, cannot correct two frictions: moral hazard and imperfect competition.

Corollary 4. *There is no contract $\phi_i = \phi$ $\forall i$ that implements the price and allocation of a frictionless market, in which both banks and households have access to and compete in an auction that induces truthful bidding, that is, avoids bid shading.*

Model with capital constraints. So far, the intermediation friction came from moral hazard. Now we add capital constraints. Suppose that each trading desk purchases $a_i(p)$ of the asset if the asset market clears at price p, and makes some loans L to an un-modeled sector of the economy, which we normalize to 0 w.l.o.g. The desk is subject to a Basel III-type capital constraint (here without risk weights): $\kappa p a_i(p) \leq E$, where E denotes the total equity capital.

From He and Krishnamurthy (2018) we know that the capital constraint binds only if the moral hazard incentive constraint binds. Given this, it is not surprising, that the auction clearing price is analogous to the price of Proposition 1 (*ii*), where bidders face a similar capital constraint.

Proposition 5. *In equilibrium $\phi_i = \frac{1}{1+m}$ for all i, and the market clears at*

$$p^* = \frac{1}{1+\lambda\kappa}\left(\mu - \left(\frac{\sigma\rho_I}{1+m}\right)\left(\frac{N-1}{N-2}\right)\frac{A}{N}\right) \tag{16}$$

with $\lambda = \frac{A}{NE}\left(\mu - \phi\sigma\rho_I\left(\frac{N-1}{N-2}\right)\frac{A}{N}\right) - \frac{1}{\kappa}$ when $\kappa p^* \frac{A}{N} \neq E$ and $\lambda = 0$ otherwise.

Trading desk i buys amount $\frac{A}{N}$, its bank obtains $\phi\frac{A}{N}$, and each mass of households $(1-\phi)\frac{A}{N}$.

Does competition matter? We now analyze whether imperfect competition in the asset market matters for whether and how the asset price is affected by intermediary frictions.

To vary the degree of competition, we vary the number of intermediaries (or trading desks) who compete for the asset. More bidders in an auction translates into higher competition. Note that this is different to the main text, in which we measure competition by the extent to which the market price that arises in the market with market power differs to the price that would arise in a perfectly competitive market. Measuring this price wedge directly gives a more precise idea of the impact of market power on prices than counting the number of market participants. However, since this wedge is endogenous, it is less useful for analyzing how changes in market power affect prices. Crucially, both measures of competition are qualitatively identical in that the price wedge (and the price impact) decreases monotonically when the number of market participants increases. In both cases, we say that competition increases.

Corollary 5. *Define $\bar{N} : \mu = \frac{(4+\bar{N}(2\bar{N}-5))\phi\rho_I\sigma A}{(\bar{N}-2)^2\bar{N}}$ and let $\mu > 0$.*

(i) Intermediary financing frictions always affect the asset price.

(ii) Let the asset market become less competitive in that N decreases to N'. For $N' \geq \bar{N}$, the shadow cost of the capital constraint increases, so that market price is more strongly affected by the capital constraint. For $N' < \bar{N}$, the shadow cost decreases and the market price is less strongly affected by the capital constraint.

Competition matters in two ways. First, with imperfectly competitive asset markets, it is no longer the case that intermediation frictions—either moral hazard or capital constraints—can be corrected by choosing intermediary remuneration, ϕ_i, optimally. To overcome (or at least reduce) the extra friction which arises from the fact that the asset market isn't perfectly competitive, a more complex remuneration scheme would be necessary.

Second, when the market is less competitive as a result of fewer intermediaries competing for the asset, the shadow cost of the capital constraint changes. Intuitively, a positive shadow cost guarantees that the capital constraint binds: $\kappa p^* a_i(p^*) \leq E$. The shadow cost is higher, the larger $p^* a_i(p^*)$ would be relative to E in a setting without the constraint. Thus to understand how the shadow cost changes, we must think through how $p^* a_i(p^*)$ changes as the auction becomes less competitive, i.e., the number of bidders N decreases. There are two opposing effects. On the one hand, each bidder wins more: $a_i(p^*) = \frac{A}{N}$ increases. On the other hand, the less competitive auction clears at a lower price: p^* decreases. When the quantity effect dominates the price effect, the shadow cost increases as N decreases. Whether this is true or not depends on how competitive the market is, i.e., the number of bidders. If the degree of competition is sufficiently strong ($N' \geq \bar{N}$), the quantity effect dominates, otherwise ($N' < \bar{N}$) the price effect dominates.

C Proofs

Proof of Proposition 1. Consider dealer i who takes the behavior of all other dealers as given and determine her best-response. Given CARA utility and $R \sim N(\mu, \sigma)$, the dealer maximizes $\mu[a_i + z_i] - \frac{\rho_i \sigma}{2}[a_i + z_i]^2 - a_i p + \lambda_i [E_i - \kappa_i p(a_i + z_i)]$ with respect to a_i for a given p subject to market clearing. The point-wise first order condition w.r.t. a_i for each clearing price p is

$$\mu - \rho_i \sigma [a_i + z_i] = (1 + \lambda_i \kappa_i)(p + \Lambda_i a_i) \Leftrightarrow a_i(p) = ((1 + \lambda_i \kappa_i)\Lambda_i + \sigma \rho_i)^{-1}(\mu - \sigma \rho_i z_i - (1 + \lambda_i \kappa_i)p), \tag{17}$$

where $\Lambda_i = \frac{\partial p}{\partial a_i}$ is the price impact. The necessary condition is sufficient as long as $-\rho_i\sigma - (1+\lambda_i\kappa_i)\Lambda_i < 0$, which holds in equilibrium.

In equilibrium, dealer i's price impact equals the slope of the inverse residual supply curve, $\Lambda_i = -\left(\sum_{j\neq i}\frac{\partial a_j(p)}{\partial p}\right)^{-1}$, and all dealers submit best-responses to each other. Therefore, given $a_j(p) = ((1+\lambda_j\kappa_j)\Lambda_j + \sigma\rho_j)^{-1}(\mu - \sigma\rho_j z_j - (1+\lambda_j\kappa_j)p)$,

$$\Lambda_i = \left(\sum_{j\neq i}((1+\lambda_j\kappa_j)\Lambda_j + \sigma\rho_j)^{-1}(1+\lambda_j\kappa_j)\right)^{-1} = \left(\sum_{j\neq i}(\Lambda_j + \frac{\sigma\rho_j}{1+\lambda_j\kappa_j})^{-1}\right)^{-1}. \quad (18)$$

By Proposition 1 in Malamud and Rostek (2017), the unique fixed point is the Λ_i that we specify in Proposition 1. The Lagrange multipliers are such that $\lambda_i(E_i - \kappa_i p^*[a_i(p^*) + z_i]) = 0\ \forall i$, where

$$p^* = \frac{\sum_i((1+\lambda_i\kappa_i)\Lambda_i + \sigma\rho_i)^{-1}(\mu - \sigma\rho_i z_i) - A}{\sum_i((1+\lambda_i\kappa_i)\Lambda_i + \sigma\rho_i)^{-1}(1+\lambda_i\kappa_i)}. \quad (19)$$

This equilibrium exists if $\lambda_i \geq 0$ for all i.

When dealers are identical ($\rho_i = \rho, z_i = z, E_i = E, \kappa_i = \kappa$), the Lagrange multipliers are all the same, and the demand function simplifies to equation (4) given that $\Lambda_i = \frac{\sigma\rho}{N-2}\frac{1}{1+\lambda\kappa}$. The market clears at $p^* = \frac{\mu(N-2)N - \rho\sigma((N-1)A+(N-2)Nz)}{(1+\kappa\lambda)(N-2)N}$. The capital constraint $E = \kappa p^* \frac{A}{N}$ characterizes λ if $\lambda > 0$: $\lambda = \frac{A}{NE}\left(\mu - \rho\sigma z - \sigma\rho\left(\frac{N-1}{N-2}\right)\frac{A}{N}\right) - \frac{1}{\kappa}$. □

Proof of Proposition 2. We derive the equilibrium by guess and verify. For this, take the perspective of dealer i and assume all others play the equilibrium of the proposition. Call the guessed equilibrium demand function $a_i^G(\cdot)$. We want to show that it is optimal for dealer i to respond with the same function.

The dealer anticipates that she will hold $\eta(p)$ in asset exposure at the end of the game if the auction clears at price p. She maximizes $\mu(a_i+z_i) - \frac{\rho\sigma}{2}(a_i+z_i)^2 - a_i p + \lambda_i(E - \kappa p\eta(p))$ point-wise w.r.t. a_i for each p subject to market clearing. The rearranged point-wise necessary condition is

$$a_i(p) = (\Lambda + \sigma\rho)^{-1}(\mu - \alpha\kappa\Lambda\lambda_i\kappa - \sigma\rho z_i - (1 - \beta\lambda_i\kappa\Lambda)p), \tag{20}$$

with Λ being the slope of the inverse residual supply curve. The condition is sufficient and therefore constitutes a best-response as long as $-(\rho\sigma + \Lambda) > 0$. This holds by assumption.

If dealer i submits $a_i(\cdot)$ and all other dealers $j \neq i$ submit $a_j^G(\cdot)$, the market clears at price $p^{BR} : a_i(p^{BR}) + \sum_{j\neq i} a_j^G(p^{BR}) = A$. The capital constraint $E - \kappa p^{BR}\eta(p^{BR}) = 0$ pins down a mapping between λ_i and λ that we call $\lambda_i(\lambda)$. The guessed equilibrium exists if there is a fixed point $\lambda_i(\lambda) = \lambda$, such that $\lambda \geq 0$. Furthermore, Λ, which solves $\Lambda = -\left(\sum_{j\neq i} \frac{\partial a_j^G(p)}{\partial p}\right)^{-1}$ must be strictly positive.

Given that $\lambda_i(\cdot)$ is continuous and maps from a closed interval of $\mathbb{R}$, namely $[0, \infty)$, to itself, it has a fixed point by Brouwer fixed-point theorem. Whether this fixed point is such that $\lambda \geq 0$ and $\Lambda > 0$ depends on exogenous parameters.[23] $\square$

Proof or Proposition 3. To prove this proposition, we follow Wittwer (2018) who proves existence of an equilibrium in a discriminatory price auction with random supply and independent private information. To apply Theorem 2 and Corollary 1, it suffices to replace $\overline{Q}$ by $\overline{A}$, ρ in Wittwer (2018)'s setting by $\frac{\rho\sigma}{1+\lambda\kappa}$, t_i by $\frac{1}{1+\lambda\kappa}(\mu - \sigma\rho z_i)$, and $\underline{t}$ by $\frac{1}{1+\lambda\kappa}(\mu - \sigma\rho\overline{z})$. Further, we must invert the bidding function so that it becomes demand function (13). $\square$

Proof of Proposition 4. To derive the equilibrium of the proposition, we guess and verify. We guess that there is a symmetric equilibrium in which all contracts are the same, $\phi_i = \phi$, and all trading desks i choose the same demand,

$$a_i(p) = a(p) = \left(\frac{N-2}{N-1}\right)\frac{1}{\rho_I\sigma\phi}(\mu - p), \tag{21}$$

[23]The equilibrium may not be unique. For instance, when $N = 3, \alpha = 5, \beta = 0.5, \mu = 10, E = 1, z_i = 0\ \forall i$, $\sigma = 1, \rho = 0.1, \kappa = 0.3, A = 1$, there are two solutions: $\lambda_1 = 0.23$ and $\lambda_2 = 2.80$. $\Lambda_1 = \{13.92, 0.10\}, \Lambda_2 = \{1.07, 0.11\}$. The demand curve is strictly downward sloping in all cases.

for each p, and level of effort, $s_i = 0$. To verify that this equilibrium exists and derive the functional form for ϕ, we begin in the auction stage. We let all trading desks other than i play the symmetric equilibrium and determine trading desk i's best-response in the auction. Then we find contract ϕ_i that trading desk i's intermediary and households choose assuming that $\phi_j = \phi$ for all $j \neq i$. The proof is complete when we have shown that the best-responses equal to the guessed equilibrium.

A trading desk with contract ϕ_i chooses her demand function $a_i(\cdot)$ and whether to exert effort or not, $s_i \in \{0, 1\}$, to maximize the expected utility she obtains from wealth

$$\omega_i(a_i(p), s_i) = \phi_i\{a_i(p)(R - p) - s_i\Delta\} + s_i b \tag{22}$$

point-wise for each p and subject to market clearing, i.e., $\sum_i a_i(p) = A$. If the trading desk exerts effort ($s_i = 0$), her wealth in period 2 is ϕ_i of the return that the asset will generate, which is $a_i(p)(R - p)$. If she shirks ($s_i = 1$), she obtains benefit b but suffers a loss which comes from the fact that the total generated wealth reduces by Δ due to shirking. Given her contract, the trading desk's loss is ϕ_i of that. Thus, the trading desk exerts effort if the benefit of doing so is larger than the cost, which is the case when

$$\phi_i\Delta \geq b \Leftrightarrow \phi_i(1 + m) \geq 1 \text{ where } m = \frac{\Delta}{b} - 1. \tag{23}$$

Following the same steps as in the proof of Proposition 1, we find that i chooses

$$a_i(p) = (\Lambda_i + \phi_i\rho_I\sigma)^{-1}(\mu - p) \text{ with } \Lambda_i = \frac{\phi\rho_I\sigma}{N - 2} \tag{24}$$

in response to all other trading desks choosing the equilibrium guess. The auction clears at

$$a_i(p^*) + (N - 2)\frac{1}{\rho_I\sigma\phi}(\mu - p^*) = A \Rightarrow p^* = \mu - \frac{(N - 1)\phi\phi_i\rho_I\sigma A}{(N - 2)(\phi + (N - 1)\phi_i)}. \tag{25}$$

Anticipating how trading desks behave in the auction, households and the intermediary choose ϕ_i for trading desk i to maximize their joint expected utility from wealth. Given CARA utility, this is equivalent to

$$\max_{\phi_i} \text{Welfare}(\phi_i) = \sum_{j \in \{H,I\}} a_j(p^*)(\mu - p^*) - \frac{1}{2}a_j^2(p^*)\rho_j\sigma \quad \text{subject to } \phi_i\Delta \geq b, \tag{26}$$

where $a_H(p) = (1-\phi_i)a_i(p), a_I(p) = \phi_i a_i(p)$, and $a_i(p)$, and p^* are given by (24) and (25), respectively. The solution to this problem pins down a mapping between ϕ_i and ϕ. In the symmetric equilibrium, ϕ_i must be chosen to be ϕ in response to all other contracts $j \neq i$ being ϕ. Depending on the size of Δ and b, or equivalently m, there are two solutions to this. For $m \geq 0$, the solution is $\phi_i = \frac{\Delta}{b}$, or equivalently, $\phi_i = \frac{1}{1+m}$. For $m < 0$, which is not the case we focus on, the solution is $\phi = 1 + \frac{\rho_H}{(N-2)(\rho_H+\rho_I)}$. Inserting this $\phi_i = \frac{1}{1+m}$ into the market clearing price completes the proof. □

Proof of Proposition 5. Since the moral hazard incentive constraint binds, so that $\phi = \frac{1}{1+m}$, whenever the capital constraint binds, the proof is analogous to the proof of Proposition 1 in combination with the proof of Proposition 4. □

Proof of Corollary 1. Let all λ_i's decrease, which happens, for instance, when κ_i's decrease. The corollary makes four separate statements, which we address in order. (*i*) To show that demand (3) becomes flatter, take the derivative w.r.t. p and then w.r.t. λ_i:

$$\frac{\partial}{\partial\lambda_i}\frac{\partial a_i(p)}{\partial p} = \frac{-\kappa_i\rho_i\sigma + (1+\lambda_i\kappa_i)^2\frac{\partial\Lambda_i}{\partial\lambda_i}}{(\rho_i\sigma + \Lambda_i + \kappa_i\lambda_i\Lambda_i)^2}. \tag{27}$$

This expression is negative since $\frac{\partial\Lambda_i}{\partial\lambda_i} \leq 0$ by statement (*iii*) and $\kappa_i\rho_i\sigma \geq 0$. Thus, when λ_i decreases, the (negative) slope increases towards 0, which means that demand becomes flatter; (*ii*) When demand curves become flatter, the market price must increases ceteris paribus; (*iii*) By Malamud and Rostek (2017) Proposition 1, the price impact Λ_i is monotonically increasing in α_i, which implies that it is decreasing in λ_i; (*iv*) By definition,

$markup = p^*(0) - p^*(\Lambda)$ with $p^*(\Lambda)$ as defined in (19). The markup increases when λ_i's decreases, because $p^*(0)$ increases more than $p^*(\Lambda)$. When dealers are identical, the markup has a closed form solution:

$$markup = p^*(0) - p^*(\Lambda) = \frac{1}{1+\lambda\kappa}\left(\frac{A}{N}\right)\left(\frac{1}{N-2}\right)\rho\sigma. \tag{28}$$

From this expression, we see that the markup decreases in λ. □

Proof of Corollary 2. Consider uniform price auctions. With identical dealers, the equilibrium is given by Proposition 1 (*ii*). The clearing price (19) simplifies to $p^* = \frac{1}{1+\lambda\kappa}\left(\frac{1}{N}\sum_i \mu_i - \frac{N-1}{N-2}\frac{A}{N}\sigma\rho\right)$ where $\mu_i = \mu - \sigma\rho z_i$. The markup is given in (28). With these expressions, we can take partial derivatives w.r.t. $\lambda\kappa$ to show that

$$\eta = \frac{\partial markup}{\partial\lambda\kappa}\frac{\lambda\kappa}{markup} = \frac{\partial p^*}{\partial\lambda\kappa}\frac{\lambda\kappa}{p^*} = \frac{1}{1+\lambda\kappa} - 1 \tag{29}$$

with slight abuse of notation. □

Extension. Now consider discriminatory price auctions with ex-ante identical dealers for which Proposition 3 characterizes the equilibrium. Adopting the same notation as before, the market clears at $p^*(\Lambda) = \frac{1}{1+\lambda\kappa}\left(\frac{1}{N}\sum_i \mu_i - \sigma\rho\left(\frac{N-1}{N(1-\xi)-1}\right)\frac{A}{N}\right)$ with μ_i defined as in Proposition 3 when dealers have market power, and at $p^*(0) = \frac{1}{1+\lambda\kappa}(\mu + \frac{\rho\sigma}{\xi-1}(\frac{A}{N} + \frac{1}{N}\sum_i z_i - \xi\overline{z}))$ when dealers are price-takers. The markup is

$$markup = p^*(0) - p^*(\Lambda) = \frac{1}{1+\lambda\kappa}\left(\frac{A}{N}\right)\left(\frac{\xi}{\xi-1}\right)\rho\sigma. \tag{30}$$

The remainder of the proof is identical to before. □

Proof of Corollary 3 The corollary makes three separate statements, which we address in order first for uniform price and then for discriminatory price auctions.

Consider the equilibrium in a uniform price auction of Proposition 2. (*i*) We want to show that the (negative) slope of the demand curve,

$$\frac{\partial a_i(p)}{\partial p} = -\frac{1-\beta\kappa\Lambda\lambda}{\Lambda+\sigma\rho} \text{ with } \Lambda = \frac{(N-2) \pm \sqrt{(N-2)^2 - 4\beta\kappa\lambda(N-1)\rho\sigma}}{2\beta\kappa\lambda(N-1)}, \tag{31}$$

increases (meaning that demand becomes flatter) when λ decreases. Equivalently, we want to show that

$$\frac{\partial \frac{\partial a_i(p)}{\partial p}}{\partial \lambda} = \begin{cases} -\frac{\beta\kappa}{\sqrt{(N-2)^2-4\beta\kappa\lambda(N-1)\rho\sigma}} & \text{if } \Lambda = \frac{(N-2)+\sqrt{(N-2)^2-4\beta\kappa\lambda(N-1)\rho\sigma}}{2\beta\kappa\lambda(N-1)} \\ +\frac{\beta\kappa}{\sqrt{(N-2)^2-4\beta\kappa\lambda(N-1)\rho\sigma}} & \text{if } \Lambda = \frac{(N-2)-\sqrt{(N-2)^2-4\beta\kappa\lambda(N-1)\rho\sigma}}{2\beta\kappa\lambda(N-1)} \end{cases}$$

is negative. The first case applies when $\beta > 0$, the second when $\beta < 0$. In both, the expression is negative; (ii) When demand curves become flatter, the market price must increases, ceteris paribus; (iii) By market clearing, $p^*(\Lambda) = \frac{1}{1-\beta\kappa\lambda\Lambda}\left(\mu - \alpha\kappa\lambda\Lambda - \rho\sigma\frac{1}{N}\left(\sum_i z_i - A\right)\right)$. This price together with the definition of the markup and the fact that Λ decreases in λ when $\beta > 0$ implies that the markup decreases in λ for relevant valid parameter constellations.

Consider the equilibrium in a discriminatory price auction of Proposition 3. (i) It is straightforward to see that demand (13) becomes flatter as λ decreases; (ii) As before, when demand curves become flatter, the market price must increases ceteris paribus; (iii) The markup, given in (30), decreases in λ. □

Proof of Corollary 4. To show that there is no $\phi_i = \phi$ for all i that implements the price and allocation of a frictionless market, we compute the price and allocation of such a frictionless market and compare both to the analogue in our market setting.

In the frictionless market, agent of type $j \in \{H, I\}$ submits the following demand $\bar{x}_j(p) = \frac{1}{\rho_j\sigma}(\mu - p)$. Intuitively, each agent submits the marginal utility she achieves from winning amount $a_j(p)$, conditional on the auction clearing at p. The market would clear at price $\bar{p}^* = \mu - (\rho_H + \rho_I)\sigma\frac{A}{N}$ at which $N\bar{a}_H(\bar{p}^*) + N\bar{a}_I(\bar{p}^*) = A$. Households obtains $\bar{a}_H(\bar{p}^*) = \frac{A}{N}\frac{\rho_H+\rho_I}{\rho_H}$; an intermediary obtains $\bar{a}_I(\bar{p}^*) = \frac{A}{N}\frac{\rho_H+\rho_I}{\rho_I}$.

In contrast, by Proposition 4, the actual market clears at $p^* = \mu - \rho_I \sigma \phi \left(\frac{N-1}{N-2}\right) \frac{1}{N} A$, households obtains $a_H(p^*)(1-\phi)\frac{A}{N}$, and an intermediary obtains $a_i(p^*) = \phi \frac{A}{N}$.

Comparing the prices and allocations across market settings, we see that it is not possible to obtain the frictionless price and frictionless allocation with the same ϕ. We can only obtain one of the two. $\square$

Proof of Corollary 5. Statement (i) follows from Proposition 4 and Corollary 4. To show statement (ii) we only need to determine how $\lambda > 0$ changes in N, since we already know that the market price increases when λ decreases.

$$\frac{\partial \lambda}{\partial N} = \frac{A(-\mu(N-2)^2 N + (4 + N(2N-5))\phi \rho_I \sigma A)}{E(N-2)^2 N^3}$$

$$\frac{\partial \lambda}{\partial N} \begin{cases} < 0 & \text{if } \mu > c(N) = \frac{(4+N(2N-5))\phi \rho_I \sigma A}{(N-2)^2 N} \\ > 0 & \text{otherwise} \end{cases}$$

Note that cutoff $c(N)$ strictly decreases in N, and converges to 0 as $N \to \infty$. Therefore, given $\mu > 0$, there is some $\bar{N}$ at which $\mu = c(\bar{N})$ so that for $N > \bar{N}$, $\frac{\partial \lambda}{\partial N} < 0$ and for $N < \bar{N}$, $\frac{\partial \lambda}{\partial N} > 0$. If $\mu > \frac{7}{3} \rho_I \sigma \phi A$, $\frac{\partial \lambda}{\partial N} < 0$ for any N. $\square$

D Robustness analysis

Here, we explain how we construct alternative slope measures for the event study of Section 4, and conduct robustness checks.

Alternative slope measures. First, we replace $\max_k\{a_{itsk}\}$ in (6) by the amount each dealer asked for at the highest step, k_i^*, it ever won: $a_{itsk_i^*}$. Second, we cap $\max_k\{a_{itsk}\}$ by the largest amount a dealer ever won (in % of supply). Third, we compute the local slope around the point at which the market clears. For this, we first find the market clearing yield of an auction, $yield_{ts}^C$. Then, we leverage the fact that, according to the auction rules, bid yields cannot have more than three decimal places. We determine the yield that lies just above, $yield_{ts}^{UB} = yield_{ts}^{C} + 0.001$, and the yield that lies just below the clearing yield: $yield_{ts}^{LB} = yield_{ts}^{C} - 0.001$. With this, we compute the local slope as the difference in the demand of a dealer at these two cutoff yields over the difference in the cutoff yields.[24]

Among all slope measures, measure (6) is our preferred one. It is intuitive and can be computed for essentially all demand curves for each dealer and auction. The measure can account for differences in capitalization, capital requirements and market conditions over time and across all dealers. In contrast, our alternative slope measures rely on a more restricted sample of bids. This is especially true when measuring the slope of demand locally, around the market clearing price. This local measure only considers 69% of all demand curves—those of bidders who win at auction. This is problematic, for instance, if dealer capitalization affects how aggressive the dealer bids, which is what our model predicts.

Robustness. We conduct a series of robustness checks. First, we estimate regression (7) with leverage ratios LR_{iq} instead of capital thresholds to exploit variation in these ratios (see Appendix Figure A3a). Since LR_{iq} varies over time and quarters, unlike $threshold_{iq}$ which is fixed over time with few exceptions, we can include dealer fixed effects to take out time-invariant unobservable dealer characteristics. We run two specifications, one with and one without dealer fixed effects. In both cases, the estimates are qualitatively similar to those with thresholds, but the change in the slope is less sharp when the exemption period ends. One reason for this is that dealers who were most positively impacted by the LR exemptions demanded larger amounts in 2020q2 than in subsequent quarters (see Appendix Figure A4). This mechanically increases the slope measure after the first exemption quarter.

Second, we exclude dealers who hit the pre-COVID bidding limits more frequently than

[24]One other way to construct a slope measure in other settings, would be to fit straight line through each submitted demand curve. In our setting, in which the median (maximal) number of steps in a demand curve is four (seven), this would mean running OLS regressions through only a few data points.

0.25% of the time to show that the main results are not driven by dealers who go over these limits in Appendix Figure A5.

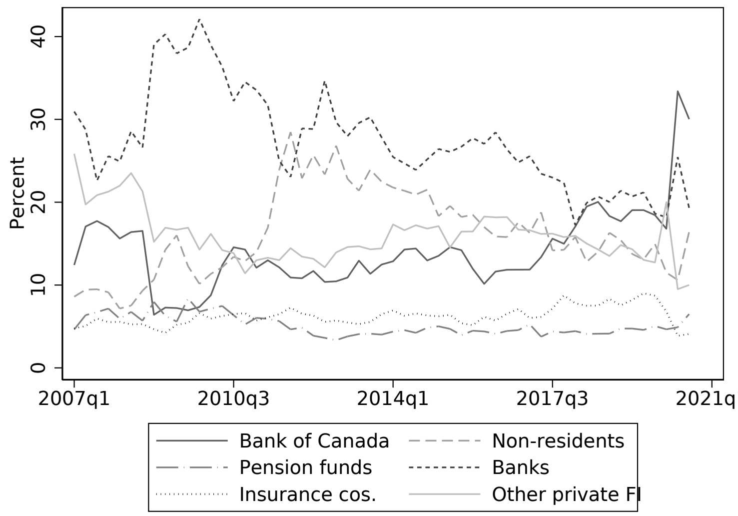
Third, we use the alternative slope measures, described above, to show that the change in the slope is not driven by the way we construct our slope measure in Appendix Figures A6. By definition, the size of the slope effect depends on how we compute the slope measure. The qualitative finding is robust across all specifications with one exception. We find no significant change in the local slope measure when relying on bank thresholds. This could be because of insufficient power given that the local measure only considers a subset of bids, or because of a bias that comes from using winning bids.

Appendix Table A1: Bid functions are approximately linear

	mean	median	sd
β_t	0.20	0.17	0.11
R_t^2	0.82	0.83	0.16
Adj. R_t^2	0.77	0.77	0.21
Within R_t^2	0.53	0.54	0.15

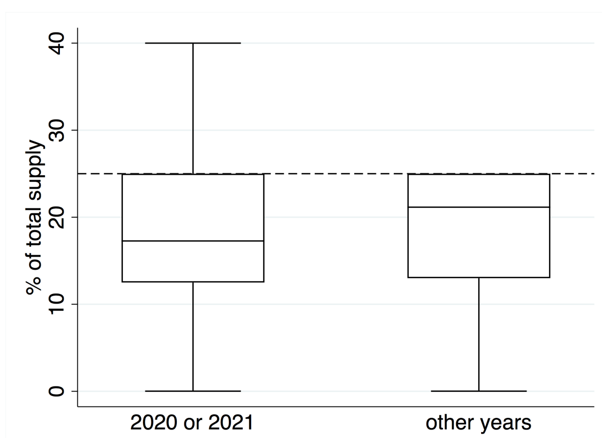
Appendix Table A1 shows the point estimate and R^2 from regressing bids on quantities in each auction: $b_{tik} = \zeta_{ti} + \beta_t a_{tik} + \epsilon_{tik}$. The subsample are bidding-functions with at least 2 steps. Bids are in yields (bps) and quantities in percentage of supply.

Appendix Figure A1: Holders of Canadian government bonds



Appendix Figure A1 shows who holds Canadian government bonds and bills from 2007 until 2021 in percentage of par value outstanding: Bank of Canada, Non-residents, Canadian pension funds, Canadian banks, Canadian insurance companies, and other private firms. The bank category holdings are mostly driven by the eight banks we focus on. They hold over 80% of the assets.

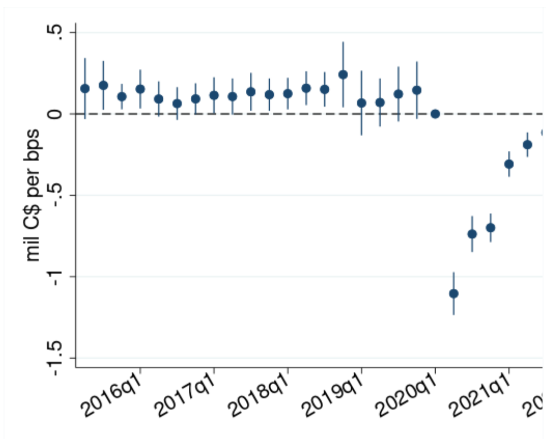
Appendix Figure A2: Maximal dealer demand as % of total supply



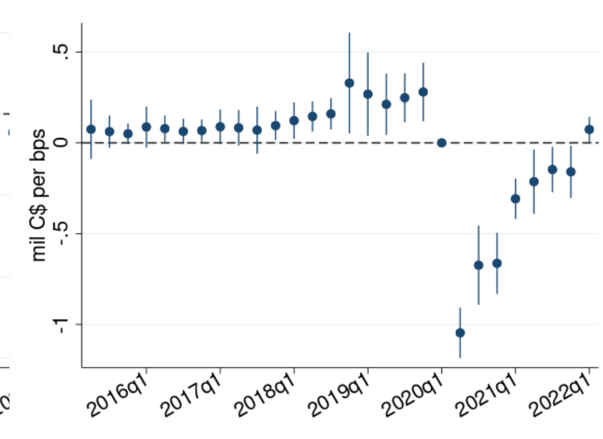
Appendix Figure A2 shows box plots of how much the average dealer demanded in an auction in 2020 or 2021 and all other years as % of auction supply. The dashed line represents the maximal bidding limit in regular times. It was increased to 40% during 2020/2021. The minimal bidding limits, which must not be met at each auction, are around 10% for most banks.

Appendix Figure A3: Change in the slope using LR

(a) Without dealer fixed effect



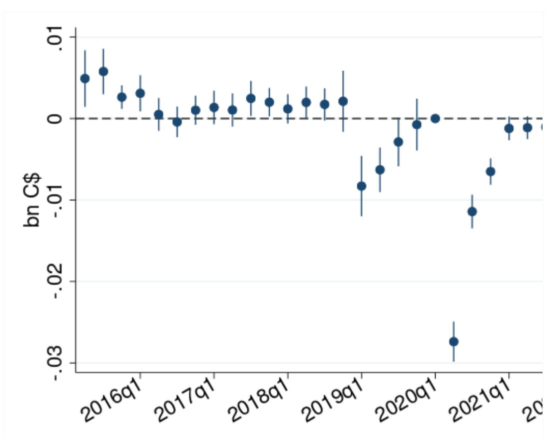
(b) With dealer fixed effect



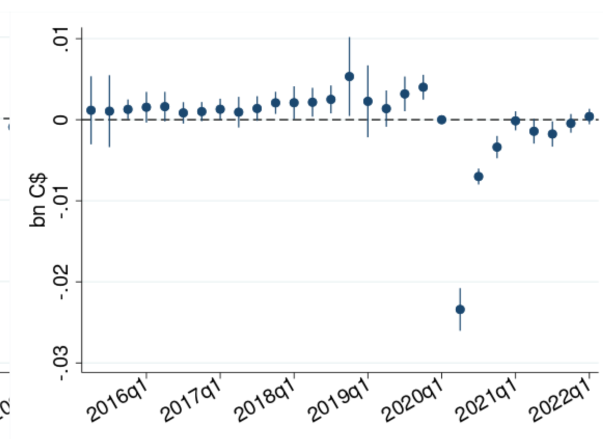
Appendix Figure A3a shows the γ_k estimates and 95% confidence intervals of the regression (7): $slope_{iqs} = \alpha + \sum_{k=-K}^{K} \gamma_k \times D_k \times LR_{ik} + \zeta_{qs} + \epsilon_{iqs}$. Figure A3b adds dealer fixed effects. All magnitudes are relative to the benchmark in 2020q1. The slopes are measured in million C$ per bps, leverage ratios are in bps. Standard errors are clustered at the dealer level in (b).

Appendix Figure A4: Event study with maximal demand, i.e., $\max_k\{a_{i,t,s,k}\}$

(a) Without dealer fixed effect



(b) With dealer fixed effect

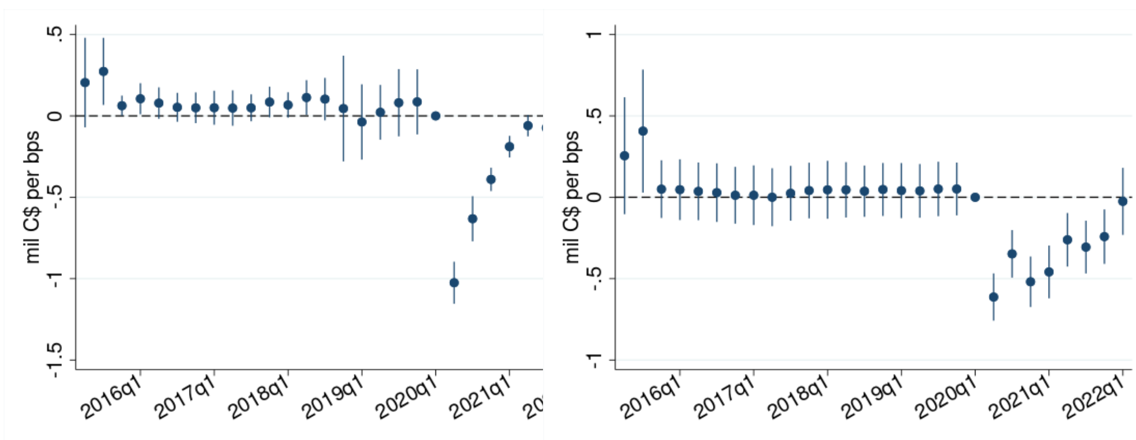


Appendix Figure A4a shows the γ_k estimates and 95% confidence intervals of the regression (7): $a\_max_{iqs} = \alpha + \sum_{k=-K}^{K} \gamma_k \times D_k \times LR_{ik} + \zeta_{qs} + \epsilon_{iqs}$ with a_max_{iqs} as average of $\max_k\{a_{itsk}\}$ per a dealer/quarter/security. In Appendix Figure A4b, we add a dealer fixed effect. Demand is measured in billion C$. Standard errors are clustered at the dealer level in (b).

Appendix Figure A5: Event study with unconstrained dealers (w.r.t. bidding limits)

(a) With LR

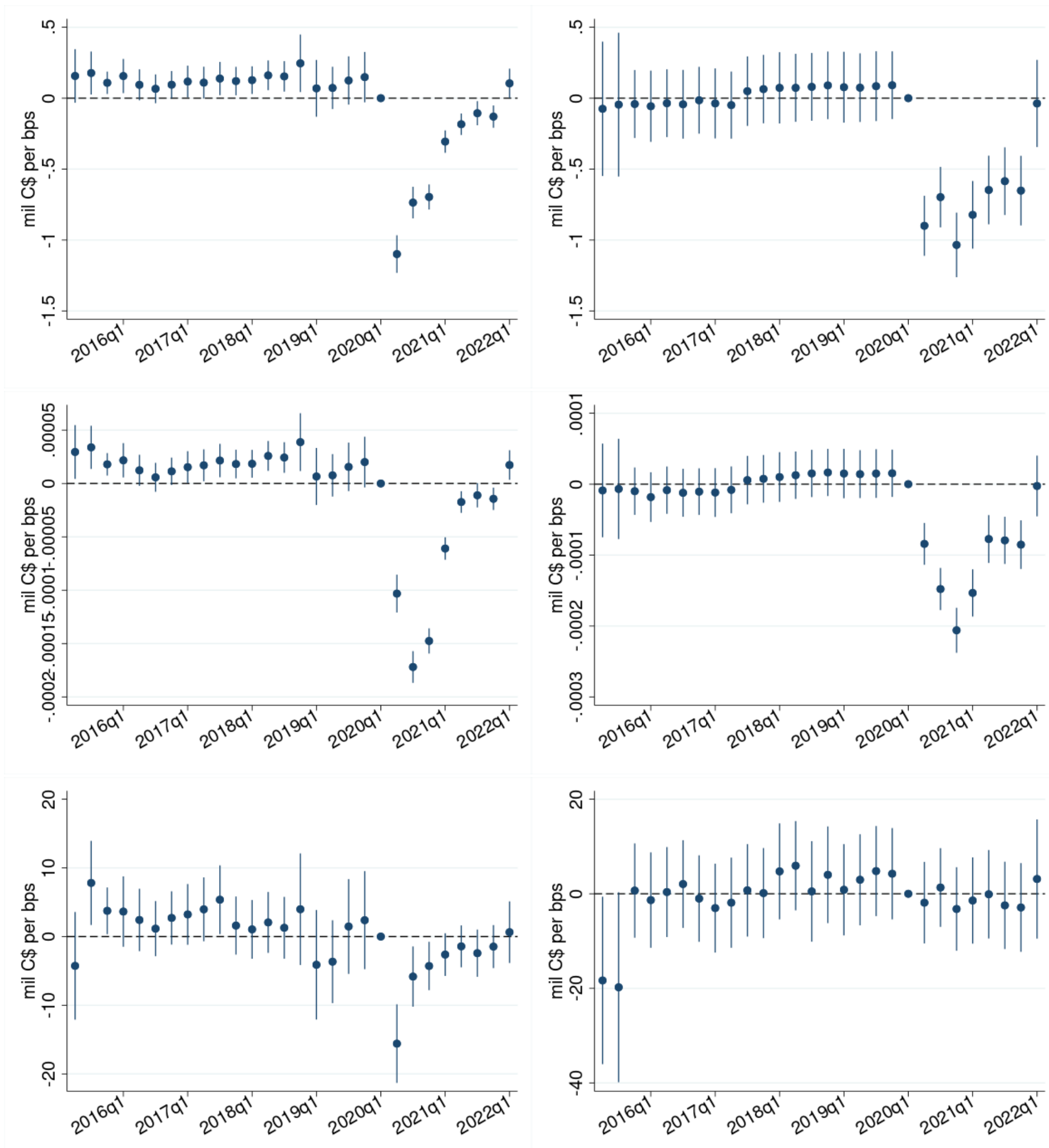
(b) With supervisory thresholds



Appendix Figure A5a shows the γ_k estimates and 95% confidence intervals of the regression (7): $slope_{iqs} = \alpha + \sum_{k=-K}^{K} \gamma_k \times D_k \times LR_{ik} + \zeta_{qs} + \epsilon_{iqs}$. Appendix Figure A5b replaces LR_{ik} by the institution-specific Basel III threshold. All magnitudes are relative to the benchmark in 2020q1. The slopes are measured in million C$ per bps, leverage ratios and thresholds are in bps.

Appendix Figure A6: Event study with alternative slope measures

(a) Using LR (b) Using supervisory thresholds



Appendix Figure A6 shows the γ_k estimates and 95% confidence intervals of the regression (7): $slope^{other}_{iqs} = \alpha + \sum_{k=-K}^{K} \gamma_k \times D_k \times LR_{ik} + \zeta_{qs} + \epsilon_{iqs}$ in (a) and with supervisory LR thresholds instead of LRs in (b) for alternative slope measures. In the first row, we replace $\max_k\{a_{itsk}\}$ in (6) by the demand at the dealer asked for at the highest step it ever won. In the second row, we cap this amount by the largest amount ever won (in % of supply). In the third row, we use a local slope measure, defined in Section 4. All magnitudes are relative to the benchmark in 2020q1. The slopes are measured in million C$ per bps, leverage ratios and thresholds are in bps. All graphs look similar when we include dealer fixed effects.