From Market Making to Matchmaking: Does Bank Regulation Harm Market Liquidity?

Gideon Saar (Cornell Johnson)
Jian Sun (MIT Sloan)
Ron Yang (Harvard Business School)
Haoxiang Zhu (MIT Sloan and NBER)

Microstructure Exchange

Post-(the last) crisis bank regulations

- Basel 2.5 increased capital requirement for the trading book.
- Basel III strengthened risk-based capital requirement and leverage ratio requirement
  - G-SIBs face tighter constraints
- Basel III also introduced liquidity requirement, including liquidity coverage ratio (LCR) and net stable funding ratio (NSFR)
- Volcker rule (US) prohibits proprietary trading, and therefore impose the need to differentiate market making from prop trading.
- Total loss absorbing capacity (TLAC), stress test, etc.

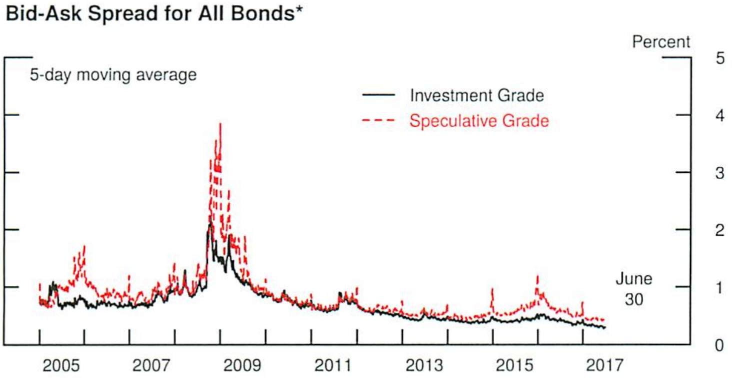
$\Longrightarrow$ Concerns about market liquidity

Bank for International Settlement, Committee on the Global Financial System (2016):

> *Risk-taking by market-makers can add to market liquidity by providing depth. At the same time, it can erode market-makers' own resilience if it is not supported by adequate capital and robust risk management practices. Thus, resilience comes at a cost, and experience suggests that the pre-crisis price of immediacy did not reflect this cost. Underpriced liquidity services were predicated on expectations of an implicit public sector backstop for major financial institutions. In that setup, the key market-makers represented a source of illiquidity contagion. Post-crisis regulatory reform aims at addressing these weaknesses...*
>
> *Improved resilience of market-makers, brought about by regulation, raises the cost of market intermediation.*

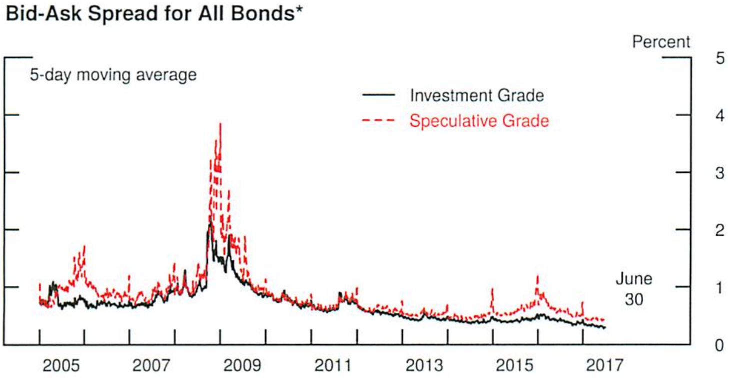
However, existing evidence on liquidity is mixed and nuanced

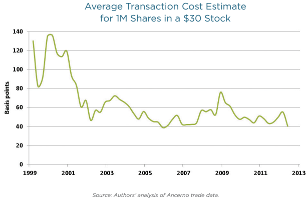
Source: Federal Reserve Board Corporate Bond Liquidity Reports (Q2 2017). Bid-ask spread is trade size-weighted average dealer bid prices and ask prices. Excludes 144a bonds.



However, existing evidence on liquidity is mixed and nuanced

Source: Federal Reserve Board Corporate Bond Liquidity Reports (Q2 2017). Bid-ask spread is trade size-weighted average dealer bid prices and ask prices. Excludes 144a bonds.





Empirical findings in the corporate bond markets:

- Average cost of trade is lower than (at least comparable to) pre-crisis levels.
- The cost of immediacy (taking liquidity from dealers) goes up.
- Market making and capital commitments went down; agency trading went up.
- Bank dealers retrenched; non-bank dealers stepped up (insufficiently).

Empirical findings in the corporate bond markets:

- Average cost of trade is lower than (at least comparable to) pre-crisis levels.
- The cost of immediacy (taking liquidity from dealers) goes up.
- Market making and capital commitments went down; agency trading went up.
- Bank dealers retrenched; non-bank dealers stepped up (insufficiently).

Not yet clear:

- How does bank regulation affect investor welfare?
- Transaction-based measures cannot capture (i) lost trading opportunities or (ii) costly delays.
- What is the appropriate policy response, if any?

Empirical findings in the corporate bond markets:

- Average cost of trade is lower than (at least comparable to) pre-crisis levels.
- The cost of immediacy (taking liquidity from dealers) goes up.
- Market making and capital commitments went down; agency trading went up.
- Bank dealers retrenched; non-bank dealers stepped up (insufficiently).

Not yet clear:

- How does bank regulation affect investor welfare?
- Transaction-based measures cannot capture (i) lost trading opportunities or (ii) costly delays.
- What is the appropriate policy response, if any?

We need a model!

Main idea of this paper

Premise:

- Bank dealers provide liquidity in two ways:
  - Market making (or principal trading), which primarily relies on balance sheet
  - Matchmaking (or agency trading), which primarily relies on (search) technology
- Bank dealers jointly optimize the use of balance sheet and technology.
- Bank dealers have market power (not fully competitive).

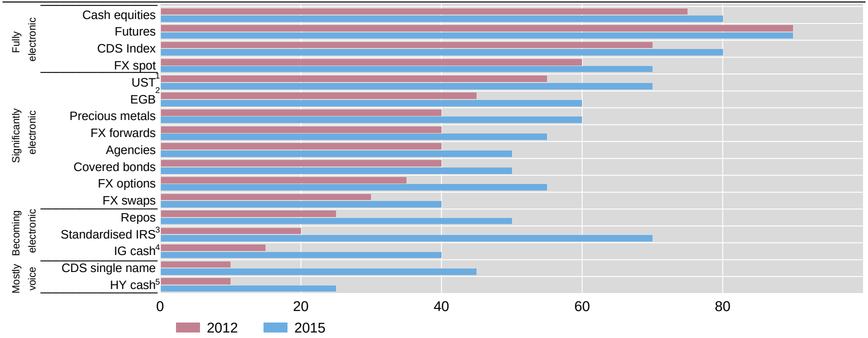
Consequence:

- Due to market power, bank dealers do too little matchmaking to preserve market making profits, although investors would prefer matching.
- Bank regulation increases the balance sheet cost of bank dealers and encourage them to do more matchmaking.
- Potential competition from non-bank dealers further prompts this shift to matchmaking.
- During the transition to matchmaking, an increase in banks' balance sheet cost generally **improves** investor welfare.

The shift toward technology is almost universal

State of electronification in various asset classes

In per cent Graph 3



[1] US Treasuries. [2] European government bonds. [3] Standardised interest rate swaps. [4] Investment grade cash bonds. [5] High-yield cash bonds.

Sources: Greenwich Associates (2014); McKinsey & Company and Greenwich Associates (2013).

Source: BIS (2016)

Details of empirical evidence on US corporate bond markets

Average transaction costs (price-based measures): improvement or no change.

- Mizrach (2015): Spreads and price impact of trades dropped after the crisis to below pre-crisis level.
- Adrian, Fleming, Shachar, and Vogt (2017): Spreads and price impact declined.
- Anderson and Stulz (2017): Price impact and spreads marginally better after regulation.
- Bessembinder, Jacobsen, Maxwell, and Venkataraman (2018): average customer execution costs have not increased after regulations were imposed.
- Trebbi and Xiao (2019): No evidence for deteriorating liquidity around post-crisis regulations.

Cost of immediacy: increase.

- Bao, O'Hara, and Zhou (2018): (downgrades of bonds to junk status) cost of immediacy increased after implementation of Volcker Rule.
- Dick-Nielsen and Rossi (2018): (exclusions from the Barclays Capital Investment-grade bond index) cost of immediacy increased after financial crisis.
- Choi and Huh (2017): trading costs for unmatched (i.e., market making) trades increased in the post-regulation period, and the increase is driven by bank dealers.

Share of trading mechanisms: Matchmaking increases; market making goes down.

- Matchmaking has increased following the crisis and the implementation of post-crisis regulations. Driven by bank dealers. (Bao, O'Hara, and Zhou (2018), Choi and Huh (2017), Schultz (2017)).
- Bank dealers committing less money to market making (Bao, O'Hara, and Zhou (2018)).
- Non-bank dealers increased capital commitment to market making and amount of principal trading (insufficient to offset bank dealer decrease), but decreased matchmaking (Bao, O'Hara, and Zhou (2018), Bessembinder, Jacobsen, Maxwell, and Venkataraman (2018)).

Volume:

- Overall trading volume and bond issuance increased. Turnover decreased (increased) in more (less) active bonds (BIS CGFS (204), Mizrach (2015), Adrian et a. (2017)).

Theoretical literature

- An, Song, and Zhang (2017); An and Zheng (2017); Li and Li (2017); Cimon and Garriott (2019)
- Mussa and Rosen (1978); Katz (1984); Johnson and Myatt (2003); Nocke and Schutz (2018); Weyl and Fabinger (2013)

Model

- One asset with expected value v (common knowledge). Everyone is risk neutral. Discount rate is r. Time $t \in (0, \infty)$.
- Infinitesimal buyers arrive as a flow with rate μ. A buyer wishes to buy one unit, with private value $x \in [0, \infty)$ with cdf G. Sellers have the same arrival rate μ and distribution G of private values. All customers are price-takers.
- A customer's private value is not observable to anyone else.

Assumption

$\zeta(x) \equiv \frac{1-G(x)}{G'(x)}$ *is non-increasing in* x*.*

- It implies that virtual valuation $\phi(x) \equiv x - \zeta(x)$ is increasing in x.

- A representative strategic bank dealer provides two services.
- Market making: Immediately takes the other side, at a spread S_B and balance sheet cost c_B.
- Matchmaking: By incurring cost I, the bank dealer matches customers to the other side with intensity H, i.e., exponential time with mean $1/H$. The bank dealer charges a fee f.
- A representative strategic non-bank dealer only does market making at spread S_{NB} and cost c_{NB}. (The non-bank dealer does not have a client base for matchmaking.)
- Our interpretation of c_B vs c_{NB}:

$$c_B = \underbrace{\text{Activity-Based Cost of Capital}}_{c_{NB}?} - \text{Implicit Subsidy} + \text{Regulatory Costs.}$$

Strategy of customers

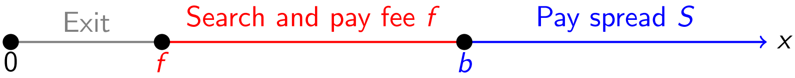
A customer's profit if his private value is x:

$$\begin{array}{rl} x - S & \text{pay the market making spread } S = \min(S_B, S_{NB}) \text{ immediately} \\ (x - f)\underbrace{E[e^{-rT}]}_{\frac{H}{H+r} \equiv \mathcal{H}} & \text{match and pay fee } f \text{ at matching time } T \sim Exp(H) \\ 0 & \text{exit the market} \end{array}$$

- An investor with private trading benefit b is indifferent between paying the spread immediately and searching:

$$b - S = (b - f)\mathcal{H}, \quad b = \frac{S - f\mathcal{H}}{1 - \mathcal{H}}. \tag{1}$$

- An investor with private trading benefit f is indifferent between searching and exiting.



Total customer welfare

Given the two thresholds f and b, the overall welfare of customers aggregated across the three ranges of x is:

$$\pi_c = \frac{2\mu}{r}\left[\underbrace{\int_{x=0}^{f} 0 \cdot dG(x)}_{\text{no trade}} + \underbrace{\int_{x=f}^{b} (x-f)\mathcal{H} dG(x)}_{\text{matchmaking}} + \underbrace{\int_{x=b}^{\infty} (x-S) dG(x)}_{\text{market making}}\right]. \tag{2}$$

Bank and non-bank dealers' problems

The bank dealer's profit is comprised of three components:

$$\pi_B = \frac{2\mu}{r}\left[\underbrace{(\mathcal{H}f - I)(G(b) - G(f))}_{\text{matchmaking profits}} + \underbrace{(S - c_B)(1 - G(b))\mathbb{I}_{S=S_B}}_{\text{market making profits}}\right], \tag{3}$$

where $\mathbb{I}_{S=S_B}$ takes the value 1 if $S = S_B$ (equivalently, $S_B \leq S_{NB}$) and 0 otherwise.

The market making profit of the non-bank dealer can be expressed as:

$$\pi_{NB} = \frac{2\mu}{r}\left[(S - c_{NB})(1 - G(b))\mathbb{I}_{S=S_{NB}}\right], \tag{4}$$

where $\mathbb{I}_{S=S_{NB}}$ takes the value 1 if $S_{NB} < S_B$ and 0 otherwise.

Equilibrium definition

An equilibrium consists of:

1. The bank dealer's choices of market making spread S_B and matching fee f;
2. The non-bank dealer's choice of market making spread S_{NB}; and
3. Each arriving customer's choice between market making (with one of the dealers), matchmaking, and refraining from trade altogether;

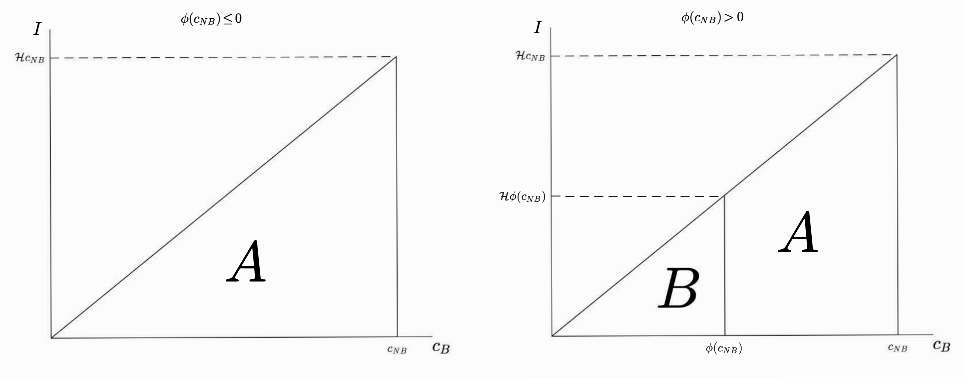
such that dealers and customers maximize expected profits.

To model a higher balance sheet cost of banks, we increase c_B and hold all else fixed.

Does bank regulation reduce liquidity and investor welfare?

Regions of equilibrium

- We focus on parameter regions in which matchmaking exists, or $I < \mathcal{H}c_B$.
- If, in addition, $c_B < c_{NB}$, there are two possible cases of equilibrium with matchmaking.
- Region A ($c_B > \phi(c_{NB})$): constrained bank dealer equilibrium, $S_B = c_{NB}$.
- Region B ($c_B < \phi(c_{NB})$): unconstrained bank dealer equilibrium, $S_B < c_{NB}$.



Intuition of the regions

Consider a monopolist bank dealer's profit-maximization problem

$$\max_{S_B \le c_{NB}} (S_B - c_B)(1 - G(S_B)). \tag{5}$$

The first-order condition is

$$c_B = S_B - \frac{1 - G(S_B)}{G'(S_B)} = \phi(S_B). \tag{6}$$

Under the technical assumption, the right-hand side is increasing.

- If $\phi(c_{NB}) > c_B$, then there exists a unique $S_B \in (0, c_B)$ such that the first-order condition holds.
- If $\phi(c_{NB}) < c_B$, then the solution is corner, $S_B^* = c_{NB}$.

Equilibrium

Proposition

When $c_B \leq c_{NB}$ and $I < \mathcal{H}c_B$, the bank dealer operates both market-making and matchmaking services, and the equilibrium is characterized as follows:

1. *If $\phi(c_{NB}) \leqslant 0$ and $I \in (0, \mathcal{H}c_B)$, there is a* constrained bank dealer *equilibrium ($S^\star = c_{NB}$), and $f^\star$ is the minimal solution of*

$$f^\star = \arg\max_f \frac{2\mu}{r}\left[(\mathcal{H}f - I)\left(G\left(\frac{c_{NB} - \mathcal{H}f}{1-\mathcal{H}}\right) - G(f)\right) + (c_{NB} - c_B)\left(1 - G\left(\frac{c_{NB} - \mathcal{H}f}{1-\mathcal{H}}\right)\right)\right]. \quad (7)$$

2. *If $\phi(c_{NB}) > 0$, then*
    1. *If $I \in (0, \mathcal{H}\phi(c_{NB}))$, there exists $c_1 \in \left(\frac{I}{\mathcal{H}}, c_{NB}\right)$, such that*
        1. *If $c_B \in \left(\frac{I}{\mathcal{H}}, c_1\right)$, there is an* unconstrained bank dealer *equilibrium ($S^\star < c_{NB}$) that satisfies the following conditions:*

$$\phi(f^\star) = \frac{I}{\mathcal{H}},\ \phi(b^\star) = \frac{c_B - I}{1-\mathcal{H}},\ S^\star = \mathcal{H}f^\star + (1-\mathcal{H})\,b^\star; \quad (8)$$

        2. *If $c_B \in [c_1, c_{NB}]$, there is a* constrained bank dealer *equilibrium ($S^\star = c_{NB}$), and $f^\star$ is the minimal solution of (7).*

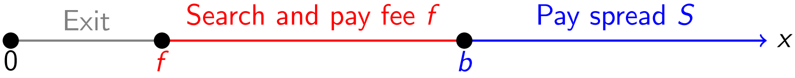
Comparative statics for constrained bank dealer equilibrium

Proposition

When c_B increases in the constrained bank dealer *equilibrium,*

1. *The spread is unchanged ($S^\star = c_{NB}$), the matchmaking fee $f^\star$ decreases, and average transaction costs decrease;*
2. *Trading volume increases, matchmaking increases, and market making decreases;*
3. *Overall customer welfare, π_c, increases.*

- Intuition for declining f: the bank dealer's profit from market making $S_B - c_B$ declines, so it also requires a lower profit from matchmaking.
- The average transaction cost is $\frac{1}{1-G(f)}[(G(b) - G(f))f + (1 - G(b))S_B]$.



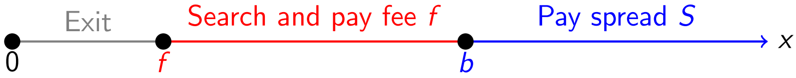
Comparative statics for unconstrained bank dealer equilibrium

Proposition

When c_B increases in the unconstrained bank dealer *equilibrium,*

1. *The spread $S^\star$ increases, the matchmaking fee $f^\star$ is unchanged, and average transaction costs increase if $c_B < (1-\mathcal{H})\,f^\star + I$ and decrease if $c_B \geqslant (1-\mathcal{H})\,f^\star + I$;*
2. *Trading volume is unchanged, matchmaking increases, and market making decreases;*
3. *Overall customer's welfare, π_c, decreases.*

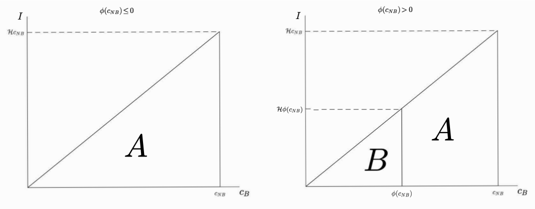
- Intuition: the unconstrained bank dealer fully passes on the higher balance sheet cost to customers. No change in per-capita profit $S_B - c_B$ and no change in matchmaking fee.
- This scenario is what commentators have in mind when they argue that regulation harms liquidity. Customer welfare is not fully measured by transaction cost.



How do we tell the two cases apart in the data?

	Spread S_B	Fee f	Avg transaction cost	Market making %	Matchmaking %	Customer welfare
Unconstrained	$\uparrow$	Flat	Hump shaped	$\downarrow$	$\uparrow$	$\downarrow$
Constrained	Flat	$\downarrow$	$\downarrow$	$\downarrow$	$\uparrow$	$\uparrow$

- The data show that average transaction costs have declined post-regulation.
- So according to the model, we are either in the constrained region A or near the right edge of the unconstrained region B.
- A further increase in c_B would increase customer welfare if we are already in region A.



The case for $c_B > c_{NB}$

Bessembinder et al (2018 JF): Bank dealers handle about 87% of principal trading (market making) post-regulation, and nonbank dealers handle 13% (up from 3% prior to 2007-08 crisis). $\Rightarrow$ c_B remains below c_{NB} but they are close.

Proposition

When $c_B > c_{NB}$ and $I < \mathcal{H} \min\{\tilde{c}_B, c_B\}$, there exists a unique equilibrium such that the non-bank dealer operates the market-making service and the bank dealer operates the matchmaking service, with $\tilde{c}_B$ as the unique solution of

$$\xi(\tilde{c}_B) - \frac{\tilde{c}_B - c_{NB}}{1 - \mathcal{H}} = 0,$$

provided that G is concave or G is convex with $G''' < 0$ and $\mathcal{H} < \frac{1}{2}$. In particular, there exists $c_2 > \max\{c_{NB}, \frac{I}{\mathcal{H}}\}$ such that,

1. *If $c_B \in (c_{NB}, c_2]$, the equilibrium is a* constrained non-bank dealer *equilibrium with $S^\star = c_B$;*
2. *If $c_B \in (c_2, \infty)$, the equilibrium is an* unconstrained non-bank dealer *equilibrium with $S^\star < c_B$.*

Comparative statics for the constrained nonbank dealer equilibrium

Proposition

When c_B *increases in the* constrained non-bank dealer *equilibrium,*

1. $S^\star = c_B$ *increases,* $f^\star$ *increases, and the change in average transaction costs is ambiguous;*
2. *Trading volume decreases, market making decreases, and the change in matchmaking is ambiguous;*
3. *Overall customer's welfare,* π_c*, decreases.*

In the (unrealistic) unconstrained nonbank dealer equilibrium, the spread charged by nonbank dealer is below c_B, so increases in c_B do not change the equilibrium outcomes.

Impact of bank regulation on market liquidity

As c_B goes up:

	Avg transaction cost	Market making provider	Customer welfare
Unconstrained bank	Hump shaped	All bank	$\downarrow$
Constrained bank	$\downarrow$	All bank	$\uparrow$
Constrained nonbank	Ambiguous	All nonbank	$\downarrow$
Data	$\downarrow$	87% bank	**?**

- Data: Customer transaction costs declined and nonbanks start to gain on market making.
- Model: We are in the constrained bank region, not too far from the constrained nonbank region.
- Customer welfare increase in bank balance sheet cost in this region!
- For investor welfare, the right action is to finish the implementation of bank regulation, rather than reverse it.
- But once nonbanks take over market making, regulation on banks should be lighter.

Regulation has always been a key driver for market structure changes

- SEC order handling rule (1997) permits investors to compete with Nasdaq dealers and requires Nasdaq dealers display their best quotes.
- SEC Reg NMS (2005) drops the protection of slow manual quotation of NYSE specialists and encourages competition among exchanges.
- FINRA TRACE brings transparency in U.S. corporate bond markets and structured products.
- These regulations are not without controversies, but on balance they all improved market quality and reduced investors' transaction costs.

Regulations help investors when they reduce market power of dominant players, regardless of their entity types (exchanges, dealers, or others).

Multiple heterogeneous bank competing

- The single-bank assumption in the baseline model is not critical for our key result that customer welfare increases in bank regulatory cost under certain conditions.
- The same qualitative result holds in a multi-bank variant of the model.
- There are N bank dealers and one nonbank dealer. All bank dealers have the same balance sheet cost c_B and search cost I.
- Customers have "taste" for banks. Customer i's taste for bank j is $\epsilon_{ij} \sim F(\sigma)$, where a larger σ means higher variance of taste shocks.
- Timeline:
  - Each bank j picks S_j and f_j and nonbank picks S_{NB}, all simultaneously.
  - Total customer arrival rate is $2N\mu$. Customers observe their tastes $\{\epsilon_{ij}\}$ and picks a bank to affiliate with.
  - Customer i observes her private value for trading and then the baseline model happens.
  - Customers cannot trade with unaffiliated banks.

Main steps of solving the multibank model

- We focus on region $c_B < c_{NB}$.
- Denote by $s_j(S_j, f_j, S_{-j}, f_{-j})$ bank dealer j's market share.
- Each bank dealer j solves

$$\max_{0 \le f_j \le S_j \le c_{NB}} \Pi_j \equiv \frac{2N\mu}{r} \times s_j(S_j, f_j, S_{-j}, f_{-j}) \times \pi_j, \tag{9}$$

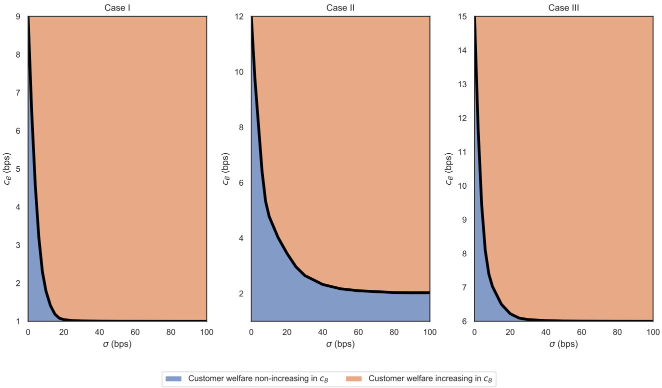
where the per-capita expected profit is

$$\pi_j = (\mathcal{H} f_j - I)(G(b) - G(f_j)) + (S_j - c_B)(1 - G(b)). \tag{10}$$

- New effect here: prices S_j and f_j also affect market share s_j.
- If ϵ_{ij} has logistic distribution with variance $\sigma^2/6$, then s_j has a closed form solution of the shape $\frac{\exp(z(S_j,f_j)/\sigma)}{\sum_k \exp(z(S_k,f_k)/\sigma)}$, where z is the customer welfare function without the taste shock.

Customer welfare can still increase in c_B as long as there is market power

- The model can be solved numerically for any integer N and $\sigma > 0$.



$N = 10$ and G is exponential with $E(G) = 10$ bps. In Case I, $c_{NB} = 9$ bps and $I/\mathcal{H} = 1$ bp. In Case II, $c_{NB} = 12$ bps and $I/\mathcal{H} = 1$ bp. In Case III, $c_{NB} = 15$ bps and $I/\mathcal{H} = 6$ bps.

An alternative welfare measure

- We believe customer welfare is the most direct metric to evaluate market performance.
- But we also consider a different welfare measure:

$$W = \pi_c + \pi_B + \pi_{NB} - \underbrace{\frac{2\mu}{r}\left(1 - G\left(b_1\right)\right)\left(c_{NB} - c_B\right)\mathbb{I}_{S=S_B}}_{\text{Too-big-to-fail subsidy to the bank dealer, if } c_B < c_{NB}} \tag{11}$$

$$= \frac{2\mu}{r}\left[\int_{x=f}^{b}(\mathcal{H}x - I)dG(x) + \int_{x=b}^{\infty}(x - c_{NB})dG(x)\right], \text{ if } c_B < c_{NB}. \tag{12}$$

Proposition

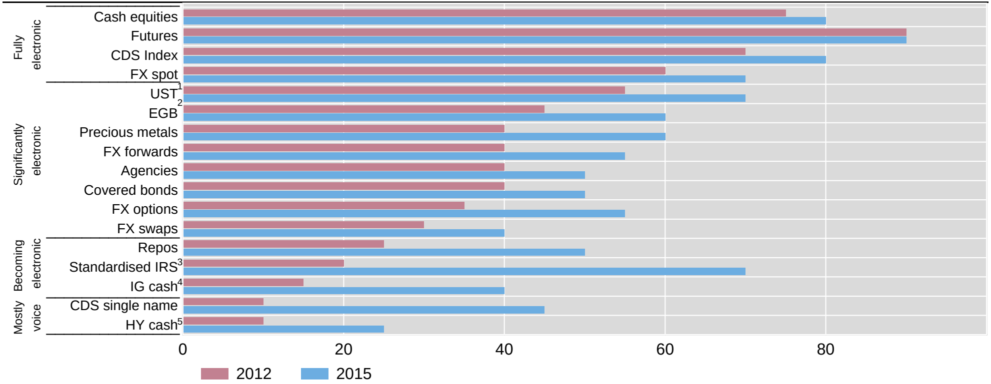
In the baseline model with $c_B < c_{NB}$ and $I < \mathcal{H}c_B$, W increases in c_B.

Conclusion

- How does bank regulation affect liquidity and investor welfare?
- The answer must take into account the change in the nature of liquidity provision from market making to matchmaking.
- Pre-crisis, technology was ready, but the transition was stalled by strategic considerations of bank dealers who held lion's share of the market with low balance sheet costs.
- Post-crisis regulation increased banks' balance sheet cost, serving as a catalyst for this healthy transition.
- In the data, customer transaction costs have declined.
- Our model implies that investor welfare has also likely increased due to regulation.
- Key result is robust to multibank competition.
- The U.S. corporate bond market is our poster boy, but similar logic applies to other fixed income markets.

State of electronification in various asset classes

In per cent | Graph 3



[1] US Treasuries. [2] European government bonds. [3] Standardised interest rate swaps. [4] Investment grade cash bonds. [5] High-yield cash bonds.

Sources: Greenwich Associates (2014); McKinsey & Company and Greenwich Associates (2013).

Source: BIS (2016)