A structural model of search and financial market liquidity*

Jamie Coen† Patrick Coen‡

October 2021

Abstract

We study financial market liquidity and its determinants. Using a rich dataset on transactions and holdings in the secondary market for sterling corporate bonds, we build and estimate a flexible model of search and trading in a decentralised asset market. We use the estimated model to study the effect of two changes in financial markets. Tighter bank capital regulation has limited impact on market liquidity, as traders adjust their holdings and search intensity to minimise its impacts. The impacts are greater in a stress, when these margins of adjustment are constrained. The introduction of trading platforms improves liquidity and welfare relative to existing bilateral trading mechanisms, but traders with low search costs who currently profit from supplying liquidity lose out.

*The views in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We are particularly grateful to Alessandro Gavazza for many helpful discussions. We are also grateful for comments by Giulia Brancaccio, Eddie Gerba, Anil Kashyap, Will Matcham, Ben Moll, Martin Pesendorfer, Mark Schankerman, Pasquale Schiraldi, Dimitri Vayanos, Nick Vause and Shengxing Zhang, as well as seminar participants at the Bank of England and London School of Economics. The authors acknowledge the financial support of the Economic and Social Research Council, the Paul Woolley Centre at LSE, the French National Research Agency and the TSE-Partnership foundation.

†London School of Economics & Bank of England, w.j.coen@lse.ac.uk.

‡Toulouse School of Economics, patrick.coen@tse-fr.eu.

1 Introduction

Key financial assets trade in decentralised markets. These markets are illiquid: it is time-consuming and costly to make a trade, and as a result trading is infrequent. Liquidity, by enabling assets to flow to those who wish to hold them, plays a key role in efficient market functioning. Illiquidity, by undermining this process, is a key contributor to market fragility and systemic risk.

Market liquidity is the result of traders' decisions, and changes to markets that affect these decisions affect liquidity. Two key changes in financial markets in recent years have focused attention on market liquidity. Firstly, tighter bank capital regulation since the financial crisis is thought to have negatively impacted liquidity in decentralised markets. Secondly, the increased prevalence of electronic trading platforms, by introducing more efficient trading mechanisms into these markets, has the potential to improve liquidity.

This paper works towards a quantitative understanding of market liquidity and its determinants. Based on a rich dataset on transactions and holdings in the secondary market for sterling corporate bonds, we build and estimate a flexible model of search and trading in a decentralised asset market. In the model, traders' search intensity, asset holdings and the terms at which they trade are endogenously determined, and traders' costs of searching vary. We assess the model's quantitative implications, and study two counterfactual scenarios: tighter bank capital regulation, and the introduction of trading platforms.

To the best of our knowledge, we are the first to take a structural approach to studying the impact of changes in banking regulation and trading technology in decentralised markets. This enables us to make the following contributions. First, we show how traders adjust their search and asset holdings in response to bank capital regulation, and show that these responses mean the costs of capital regulation in decentralised markets are limited. However, in times of stress these responses are constrained and the costs are greater. Second, we show that whilst trading platforms improve liquidity and welfare, the most efficient traders are worse off as the returns they earn from supplying liquidity are diminished. They thus have no incentive to help introduce these platforms. We are also the first to set out a model of search in decentralised markets with endogenous search and unconstrained holdings. This enables us to show how traders use search to manage their inventories and respond to shocks.

The starting point for our work is rich data on the transactions, holdings and identities of traders of sterling corporate bonds. The data are novel, relative to datasets commonly used in the literature, in two key ways: (1) they include the identities of both dealers *and customers*,[1] which means we can study how the characteristics of all traders drive their trading decisions; (2) they include data on the *holdings* of traders as well as their transactions, enabling us to study the role of traders' inventories in driving trading patterns.

We set out a number of empirical facts which inform our work. We study the frequency with which traders offset trades by buying and then selling (or selling and then buying) the same bond within a short interval. We compare the rate at which firms trade and the rate at which they offset trades, and show that the difference between them is incompatible with a model of exogenous search intensity: participants increase search intensity endogenously to offset trades. We then compare offsetting rates to inter-temporal variation in capital regulation levels and cross-sectional variation in capital regulation across bond types, and show that offsetting is more frequent when capital regulation is tighter.

We then explore which market participants intermediate. Over a quarter of trades do not involve a traditional dealer bank at all, suggesting that customers may be supplying liquidity to each other and to dealers. Traders that trade frequently are more likely to receive a spread rather than pay a spread, suggesting that in general it is the fast traders that intermediate. Finally, whilst capital regulation on banks became much more stringent over our sample period, the average spreads that dealer banks earn – a typical measure of market liquidity – did not increase.

Our data and empirical findings guide our modelling decisions in the following ways. The fact that a trader's trading frequency appears to respond to their state and the characteristics of the asset being traded leads us to consider *endogenous search intensity*. Given that trading occurs both between dealers and between customers and the customer/dealer role is not always well defined, and the fact that our data contain the trading behaviour of all types of agents, we treat *intermediation as endogenous*.

We set out and estimate a model of trading in a decentralised financial market, building on Üslü (2019) by endogenising search intensity. In the model, a continuum of forward-looking traders trade an asset with each other in a market characterised by search frictions. Traders' holdings of the asset are unrestricted. Traders face random shocks to the utility they derive from holding the asset, which creates heterogeneity in liquidity needs and thus gains from trade. Traders choose how frequently to meet a counterparty subject to a convex cost of searching, where these costs vary across traders. A trader who meets a counterparty

[1] Comparable datasets in the US, for example, do not include customer identities.

draws this counterparty randomly from the trading population. Intermediation in this model is endogenous, in the sense that rather than designating traders as intermediaries, we study how *ex ante* heterogeneity in search cost and *ex post* heterogeneity in liquidity needs across traders determines intermediation.

We estimate the model by matching theoretical moments to those we observe in the data. The novel aspects of our data, notably the presence of identifiers for all market participants, allows us to exploit the identifying information contained in a broad range of moments across and within traders. The search cost distribution is identified by the trading frequency distribution, whilst utility parameters and shock frequency are identified by the distribution of prices, quantities and asset holdings.

We use the estimated model to understand how traders' search decisions drive market outcomes, who intermediates, and why the decentralised equilibrium is inefficient. Traders increase their search intensity when their trading needs are high, which creates a U-shaped pattern of trading frequency in traders' asset holdings, with traders searching harder when they are far from a target asset holding.

Traders with low search costs intermediate more, facilitating their counterparties' trading needs and earning a spread in doing so. These can naturally be thought of as dealers, who traditionally intermediate in these markets. However, intermediation is not defined at the trader level: any trader can supply liquidity in a trade depending on how close they are to their target holdings, their search cost and the state of their counterparty. Our estimated model allows us to quantify this empirically: for any trader we can compute the proportion of trades (a) where they supply liquidity to their counterparty by enabling them to move closer to their target holding, and (b) where the reverse is true. Whilst low cost traders typically supply liquidity to high cost traders, in 30% of trades the reverse is true. This captures the nuances of liquidity provision more accurately than traditional measures such as calculating which parties pay a spread on average.

We run two types of counterfactual simulation. First, we examine the potential effects of capital regulation on dealer banks, critics of which argue increases trading costs and reduces market liquidity. To study this we simulate an increase in dealers' costs of holding assets. Dealers reduce their holdings of the asset and adjust their search behaviour, searching harder for an offsetting trade after a trade has led them to take on a large position in the asset. Dealers' spreads increase slightly, as they are more reluctant to take on large positions in the asset. Unregulated traders pick up the slack by increasing their asset holdings, causing their welfare to increase. These endogenous responses to regulation mean the net impact on liquidity and welfare is small. However, these impacts increase in periods when non-dealers are subject to stress, as they are less able to offset dealers' reduced willingness to hold the asset. In this sense, the negative effects of capital regulation are greater in periods of stress, indicating that liquidity is "fragile".

These findings rationalise a number of trends in markets in recent years, reported both in this paper and in the literature. Firstly, dealers have reduced their corporate bond inventories (Dick-Nielsen and Rossi, 2019). Secondly, dealers increasingly tightly manage their holdings, organising trades that offset incoming trade orders either instantly or soon after the trade (Schultz, 2017). Third, despite reports of worse liquidity, many aggregate measures of liquidity in normal times have not deteriorated (Adrian, Fleming, Shachar and Vogt, 2017). Fourth, liquidity during stress events appears to have deteriorated relative to before the financial crisis (Bao, O'Hara and Zhou, 2018; Dick-Nielsen and Rossi, 2019). Fifth, traders who traditionally do not operate as intermediaries can make money by supplying liquidity (Choi and Huh, 2018; BlackRock, 2015). Each of these can be rationalised as the product of regulation increasing dealers' costs of holding inventory, and the endogenous responses of traders.

We then study the impacts of technologies that decrease search costs, such as electronic many-to-many trading platforms. The bulk of corporate bond trades are organised bilaterally, between traders who communicate and bargain by phone. The same is true of the markets for derivatives and structured products. In recent years trading platforms have emerged as alternative trading mechanisms in these markets, where trades take place electronically and offered trades are posted to all participating traders rather than communicated bilaterally. These types of trading mechanisms offer clear efficiency benefits, though their progress has been relatively slow (The Economist, 2020).

We study the effects of these platforms by undertaking two counterfactuals: (a) reducing and homogenising search costs across traders, and (b) studying the outcome in a frictionless, Walrasian case. In each simulation, we find that platforms improve aggregate welfare, but reduce the welfare of the lowest-search-cost traders. We argue that this helps to explain the relatively slow adoption of these technologies: the most frequent traders – whose participation is required for a platform to be viable – would lose out under platform-based trading.

Below we set out our contribution to the literature. In Section 2 we set out our data and describe the institutional setting. In Section 3 we set out some key patterns in trading, intermediation and regulation in the sterling corporate bond market. In Section 4 we set out our model, and in Section 5 we describe how we estimate the model. In Section 6 we set out the results of our structural estimation and some key implications. In Section 7 we undertake counterfactual analyses, before concluding in Section 8.

1.1 Literature

Our work builds on and contributes to three main fields: work modelling search frictions in over-the-counter markets, empirical work on the determinants of market liquidity and structural models of financial markets.

Our paper contributes to the set of empirical studies of market liquidity and its determinants.[2] A set of papers including Adrian, Fleming, Shachar and Vogt (2017), Bessembinder, Jacobsen, Maxwell and Venkataraman (2018), Bao, O'Hara and Zhou (2018), Schultz (2017), Dick-Nielsen and Rossi (2019) and Choi and Huh (2018), seeks to understand the effect of post-crisis regulatory changes on market liquidity by comparing measures of liquidity before and the after the financial crisis. A related set of papers studies violations of no-arbitrage conditions – including covered-interest parity (Du, Tepper and Verdelhan, 2018) and the relationship between bond yields and credit default swap rates (Duffie, 2010*b*) – and relates these to post-crisis banking regulation. Both these literatures find that post-crisis regulation impacts market outcomes. We are the first to study regulation and market liquidity in a structural context, which enables us to identify the *mechanisms* by which regulation affects liquidity, and thus explain a number of recent trends in markets. A second set of papers studies the effects of technological changes on financial markets. Examples include Asquith, Covert and Pathak (2019) and Brancaccio, Li and Schürhoff (2020). We study the effect of technological changes that boost the ability of inefficient traders to contact counterparties, and explain why such changes may not be implemented.

A large theoretical literature studies search frictions in financial markets. Examples include Duffie, Gârleanu and Pedersen (2005, 2007), Hugonnier, Lester and Weill (2020) and Vayanos and Weill (2008).[3] Our model builds most closely on Üslü (2019), who sets out a

[2]See Vayanos and Wang (2013) for a survey.

[3]Other examples inlcude Afonso and Lagos (2015), Brancaccio, Li and Schürhoff (2020), Farboodi (2021), Farboodi, Jarosch and Shimer (2021), Gârleanu (2009), Gromb and Vayanos (2018), Lagos and Rocheteau (2007), Lagos and Rocheteau (2009), Liu (2020), Neklyudov (2019), Sambalaibat (2018) and Vayanos and Weill (2008).

theoretical model of endogenous intermediation between traders with different fixed trading speeds. We show that traders appear to condition their trading frequency on their state and the state of the market. We thus endogenise firms' trading speeds, which gives us a more flexible model with which to run counterfactuals and enables us to explain key trading patterns.[4] We show how to identify and estimate this class of model, and in so doing are able to provide quantitative results. To the best of our knowledge, ours is the first paper to estimate a search-and-matching model with endogenous intermediation, and to study how firms adjust trading frequency to manage their trading portfolios.

Finally, this paper is part of a growing literature specifying and estimating structural models of financial markets. Examples include Brancaccio, Li and Schürhoff (2020) in bond markets, Coen (2021) in the market for mutual funds, Benetton (2021) and Robles-Garcia (2020) in customer lending markets, Gavazza (2016) in a real asset market, and Coen and Coen (2019) in interbank markets. We draw on the methods in these papers to specify, identify and estimate our model.

2 Data and Institutional Setting

2.1 Data

We use a combination of four datasets. Our primary dataset is a database on corporate bond transactions maintained by the Financial Conduct Authority (FCA). This contains trade-level data on secondary-market trades in bonds where at least one of the firms is an FCA-regulated entity.[5] Our sample covers trading in sterling corporate bonds over the period from January 2012 to December 2017. For each trade, we see who is buying and selling the bond, the price, the quantity traded, the instrument traded, and the time of the trade. Relative to other datasets typically used in the literature, the advantage of this

[4]Liu (2020) also studies endogenous search, but in a setting with binary asset holdings, where traders are pre-assigned as dealers – who don't receive liquidity shocks – or customers. The fact that we leave holdings unconstrained means we can study how traders condition their search intensity on their asset holdings and how this impacts the distribution of their asset holdings, the fact that ours is a model of exogenous intermediation means we can study how search costs determine who intermediates, and the fact that we allow all traders to be hit by shocks enables us to study how traders use search to respond to shocks. Farboodi, Jarosch and Shimer (2021) study homogenous traders' ex ante investment in a search technology, whereas we allow a trader's choice of search intensity to depend on both their type and their state.

[5]In practice this means that almost all financial firms with a legal entity in the UK (including subsidiaries of foreign banks) appear in the dataset, including both banks and non-banks.

dataset is that it includes the identity of all traders, covering both dealers and customers.[6] This is crucial to allowing us to study endogenous intermediation, as we can observe the characteristics of traders who are not traditionally dealers.[7]

We add to this data on the bond *holdings* of seven banks, using Bank of England data, and 300 mutual funds, using data from Morningstar. The enables us to separate who holds an instrument from who trades it, and to relate trading decisions to traders' inventories of assets, for both traditional dealers and customers. This is key to studying the effect of bank capital regulation on liquidity, as this effect is thought to operate via dealers' inventories.

We match these three datasets with information on bond characteristics and primary issuance from Thomson Reuters' Eikon database.

2.2 The secondary market for sterling corporate bonds

The sterling corporate bond market is a key source of financing for both UK and non-UK firms. It has increased in importance since the financial crisis, with virtually all net financing raised by UK private non-financial firms from 2009 to 2016 coming in the form of bonds rather than bank lending (Bank of England, 2016). It is largely an over-the-counter market, with traders determining the terms of trade bilaterally, typically by phone (Anderson, Webber, Noss, Beale and Crowley-Reidy, 2015; Czech and Roberts-Sklar, 2019). Traders in the market consist of dealer banks,[8] asset managers, insurance companies and non-dealer banks. Dealer banks are a counterparty in around three-quarters of trades. 83% of trades are carried out on a principal basis, rather than on an agency basis.[9]

Table 1 gives key summary stats from our dataset. The average trading price is 109% of a bond's face value with significant variation around this figure. The median trade size

[6] Existing studies typically rely on TRACE data on US corporate bond transactions (for example Choi and Huh (2018), Trebbi and Xiao (2017) and Kargar, Lester, Lindsay, Liu, Weill and Zúñiga (2021)) or US municipal bond transactions (for example Brancaccio, Li and Schürhoff (2020) and Hugonnier, Lester and Weill (2020)). TRACE data comes in two forms: an academic version which includes anonymised dealer identifiers, and a regulatory version which reveals the dealer identities. In both cases, any customer is simply marked as 'C', meaning customers can neither be identified nor tracked through the data. The municipal bond dataset includes dealer identities, but again provides no information on the customer.

[7] Other studies using these data include Czech and Roberts-Sklar (2019), Czech, Huang, Lou and Wang (2021) and Mallaburn, Roberts-Sklar and Silvestri (2019).

[8] Where we define dealer banks as those firms that are permitted to trade in the primary market with national banks as well as inter-dealer brokers, who exist to facilitate trades between dealers.

[9] Of these agency trades, the bulk are a trading counterparty buying for a non-trading client – for example a wealth manager buying bonds on behalf of their clients. 95% of trades between trading firms are carried out on a principal basis.

is £100,000 with a tail of larger 'block trades' that skews the distribution right. Across the market there are around 1,000 firms trading each month.

The market is illiquid. Trading tends to be rare, with the median bond trading 5 times a month. There is large variation in price across bonds, and dispersion in price for the same bond: the R-squared from a regression of trading price on instrument fixed effects is 72%, with the remaining 28% reflecting within-bond price dispersion. Where a trader buys and sells a bond in the same week – a measure of the spread the trader earns – the median difference between the purchase price and the sale price is 0.13% of the bond's face value.

There is significant heterogeneity in trading activity both across instruments and across traders. Figure 1 shows the distribution of trading frequency across traders. The distribution is heavily positively skewed: a small set of firms trades very frequently, and a long tail of firms trades infrequently. This empirical fact leads us to treat traders' search costs as heterogeneous in our model, and in estimation to model these search costs as a Gamma distribution, which is well suited to capturing positively skewed distributions.

Table 1: Summary Statistics

	Mean	Std. Dev.	Median	25^{th} pctile	75^{th} pctile
Aggregate					
Price (%)	109	16	106	101	114
Trade size (£000)	475	991	100	14	400
Monthly volume (£bn)	25	5	26	22	28
Monthly traders	972	69	975	926	1,018
Instrument-level					
Issuance (£mn)	195	441	40	7	300
Trades per month	19	37	5	1	22
Number of traders	56	77	15	4	88
Trader-level					
Monthly volume (£000)	11,997	118,351	33	2	507
Instruments traded	62	199	7	2	27
Trades per instrument traded	5	35	2	1	4

Note: This table gives some key descriptive statistics from our data. Aggregate statistics are computed across all instruments and all traders. Instrument-level statistics show how issuance and trading vary across instruments. Trader-level statistics show how trading activity varies across traders. Trades per instrument shows the distribution of the ratio between a trader's total trades and the number of instruments in which they trade.

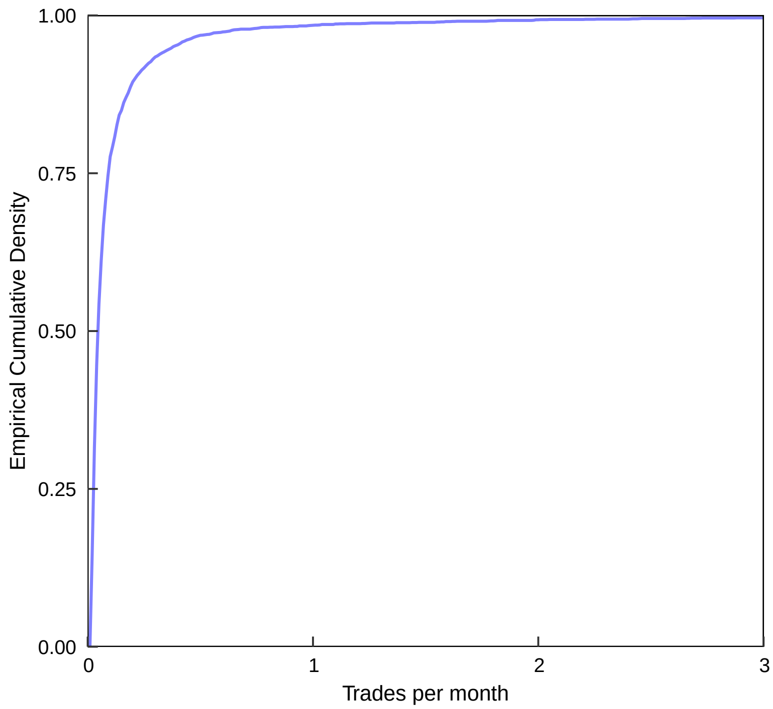


Figure 1: Trade frequency across traders

Note: This figure shows the distribution of average trading frequency per instrument across traders. For each trader we take each instrument in which they trade and compute how often they trade this instrument per month. We then average these values across each of the instruments they trade to produce an average trading frequency for each trader. This figure plots the distribution of this across traders.

2.3 Banking regulation

Banks are subject to two types of capital regulation: a leverage requirement and a risk-weighted capital requirement. The leverage requirement states that a bank's equity must exceed a given fraction of its total assets. The risk-weighted capital requirement requires that a bank's equity exceeds a given fraction of its total risk-weighted assets.[10] Risk weights vary across bonds according to their creditworthiness, and capital requirements have more than doubled since the 2008 financial crisis.[11]

[10]Where the measures of equity differ across the two requirements (BCBS, 2017).

[11]See Figure A1 for details.

3 Search, Trading and Regulation in the Sterling Corporate Bond Market

In this section we set out a number of empirical facts that motivate our questions and guide our modelling. In particular we show that patterns in trading frequencies demand a model where search intensity is endogenous, that fast traders act as intermediaries, that bank capital regulation appears to play a role in shaping trading patterns, and that the blurred distinction between liquidity demand and supply means trading is not well described by a model with exogenously determined intermediaries.

Fact 1: Trading frequency is endogenously determined

Most models of search assume traders contact each other randomly and exogenously. In particular, a trader's meeting rate does not depend on the gains to trade.[12] Here we show that the data reject this assumption.

Figure 2 shows that traders 'offset' their trades – by both buying and selling a bond in a short period of time – much more frequently than would be implied by a model where trade orders arrive exogenously. To show this, we compute the average trades per trader in each instrument in our data. We then show how often this implies a trader would both buy and sell the same bond within a 10 day period if trades arrive randomly according to a Poisson process, where the arrival rate is equal to the average trades in the instrument per firm, and a trade is equally likely to be a buy or a sell. This is the red line in Figure 2. The blue line shows the same relationship as captured in the data. The empirical pattern of trades is inconsistent with a model with exogenous search intensity. In particular, traders appear to adjust their search intensity in response to incoming trades to manage their inventory. To rationalise this, we will treat search intensity as an endogenous variable that traders choose and can condition on both their type and their state.

Fact 2: Dealer trading frequency varies with capital regulation

Capital regulation, to the extent that it increases traders' costs of holding inventory, increases the incentive for traders to minimise their asset holdings. In our second fact, we show evidence consistent with dealers doing this by increasing their search intensity after buying an asset. Figure 3 shows the likelihood that after buying an asset a dealer sells

[12] A notable exception is Liu (2020), who endogenise dealers' search intensity in a model with binary asset holdings. Farboodi, Jarosch and Shimer (2021) allow traders to make an ex ante investment to choose their meeting rate, but this meeting rate is then fixed through time.

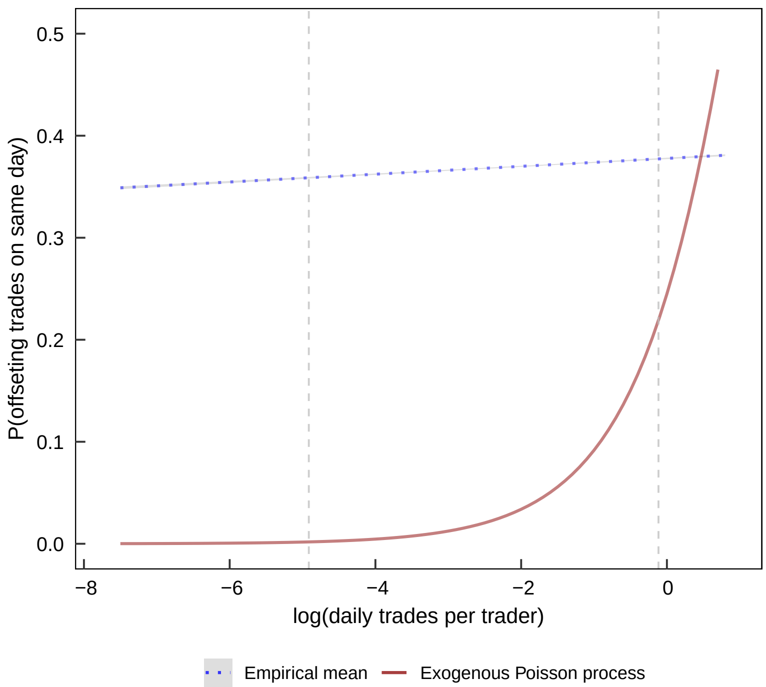
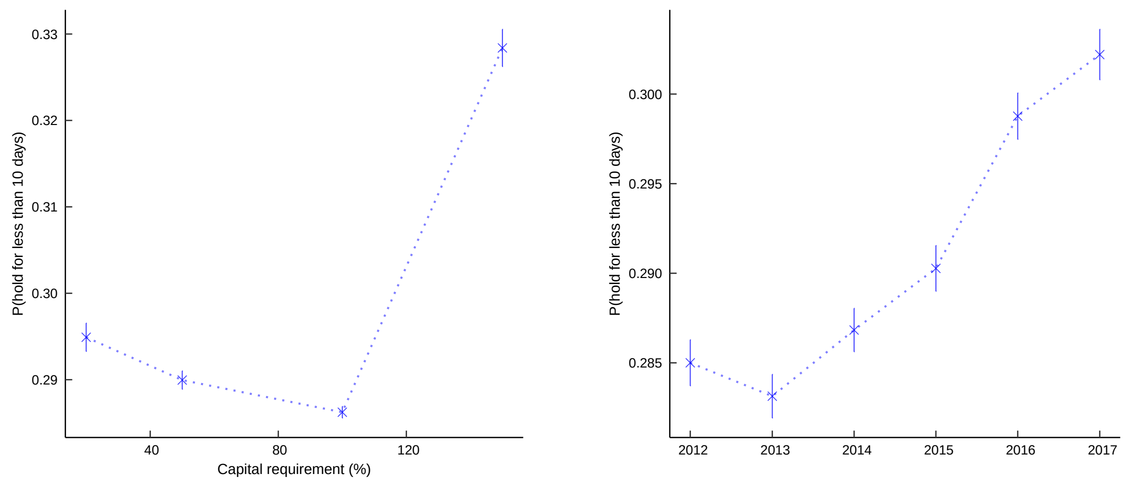


Figure 2: Trade offsetting in theory and data

Note: This figure summarises the frequency with which traders both buy and sell a bond on the same day conditional on trading in that bond, both (a) in the data and (b) as predicted by a theoretical model where buy and sell orders arrive exogenously. To compute the empirical conditional probability, we take all days on which a trader traded a bond, and create a dummy variable that equals 1 if on that day they both bought and sold that bond. We probit regress this dummy variable on the log of the average trades per trader in the instrument, and plot the fitted values in the figure. The dashed blue line shows the fitted values, and the grey shaded area around it shows the 95% confidence interval. For the theoretical conditional probability, we assume trade orders arrive randomly according to a Poisson process, where the Poisson arrival rate is equal to the average trades per trader (the log of which is plotted on the x axis). We compute the conditional probability as set out in Appendix A1. The dashed vertical lines show the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentile of the distribution of the variable on the x-axis.

the same asset within ten days, both through time and across bonds with different capital requirements. Traders were significantly more likely to offset asset purchases at the end of the sample – when capital regulation was higher – than at the start of the sample, as well as in bonds with the highest capital requirement. This suggests that the costliness of holding an asset impacts traders' trading behaviour, and this cost varies with capital regulation. This motivates our counterfactual in Section 7, where we model the impact of capital regulation on banks as an increase in dealers' costs of holding an asset.



(a) Holding period by rating

(b) Holding period through time

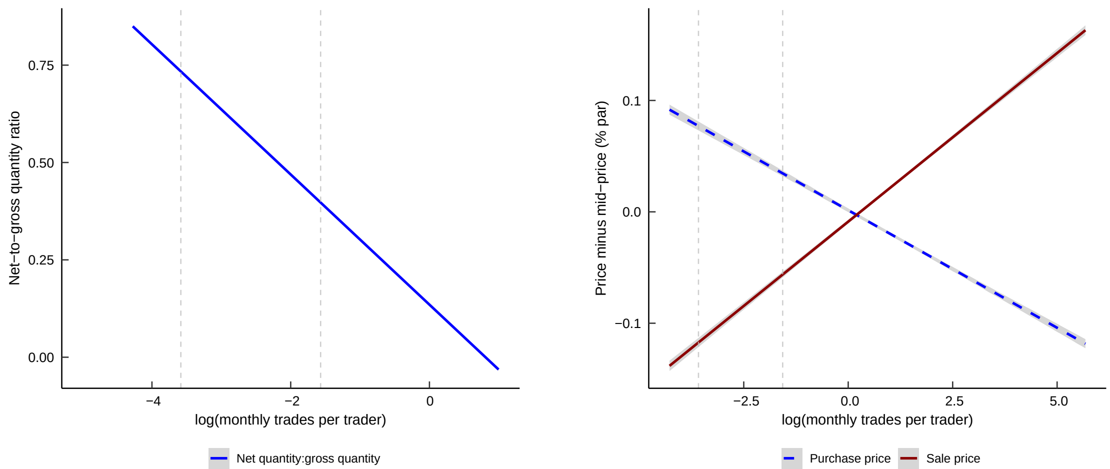
Figure 3: Patterns in asset holding periods

Note: This figure summarises how long dealers hold an asset on their balance sheet after a purchase. For each purchase by a dealer, we compute the length of time before they next sell that asset, or the asset matures. We create a dummy variable equal to 1 if this is less than 10 days, average this across all purchases and all dealers, and show how this varies by year (right panel) and by the risk weight the bond receives in banking regulation (left panel). The vertical lines show a 95% confidence interval around the point estimate. Risk weights depend on a bond's credit rating according to the standardised approach to risk weighting, as set out in BCBS (2006).

Fact 3: Fast traders intermediate

Figure 4 shows how a firm's trading speed determines whether it intermediates in bond markets or not. Conceptually, intermediaries in over-the-counter markets exhibit two key features: they trade in both directions in an instrument, buying from some traders and selling to others, and they are compensated for doing so by earning a spread. Firms that demand liquidity are more likely to follow a buy-to-hold strategy whereby they buy a bond because its value to them is high and only sell it if shocked, and pay a spread when trading. Figure 4a shows how the ratio of a firm's net trading volume to its gross trading volume depends on its trading frequency. Firms who trade a bond frequently trade in a balanced fashion –

buying and selling the bond – whilst those who trade less frequently tend to trade in one direction only. Figure 4b shows how the prices paid by traders depend on how often they trade an instrument. The purchase price is decreasing in a trader's trading frequency, whilst the sale price is increasing. As a consequence, infrequent traders typically pay a spread, whilst frequent traders earn a spread.[13] The evidence in both these figures thus suggests that frequent traders act as intermediaries in these markets. In Section 7 we show that our model replicates the patterns in these figures.



(a) Net-to-gross trading ratio

(b) Trading prices

Figure 4: Intermediation and trading speed

Note: These figures show how intermediation activity depends on firms' trading speeds. For the left figure, we take each trade by a trader, and subtract the mean trading price for that instrument in that month. We then regress this variable on the log of the average trades per month in that instrument by the firm buying the asset (blue line) and selling the asset (red line), and plot the fitted values in the figure. Firms on the left-hand side of the figure on average pay a spread, whilst those on the right-hand side earn a spread. For the right figure, we regress the absolute value of a firm's net trading in an instrument and divide it by their gross trading in the instrument. We then regress this figure – which lies between 0 and 1, on the log of the firm's average trades per month in that instrument. The dashed vertical lines show the 25$^{\text{th}}$ and the 75$^{\text{th}}$ percentile of the variable on the x-axis, and the grey shaded areas are 95% confidence intervals.

Fact 4: Average spreads have not increased over the sample period, whilst the most illiquid bonds have seen spreads fall

Figure 5 shows the distribution of bid-ask spreads earned by dealers through time. The bid-ask spread is often used as a summary measure of liquidity, as it represents the transaction cost to a trader of a 'round-trip', where they buy a bond from a dealer and then sell

[13]This is consistent with evidence of a centrality premium – whereby more central dealers earn higher spread – documented by Di Maggio, Kermani and Song (2017) in the US corporate bond market. Intuitively, Figure 4b shows a similar relationship between prices and trading frequency holds in the sterling market, and holds across all traders rather than just dealers

it back to them. The average spread has not increased. This is consistent with findings in other papers and other markets (Adrian, Fleming, Shachar and Vogt, 2017), and has been described as a puzzle given the evidence in the literature linking tighter bank regulation with diminished liquidity (Choi and Huh, 2018). The spreads on the most illiquid bonds have, in fact, decreased significantly.

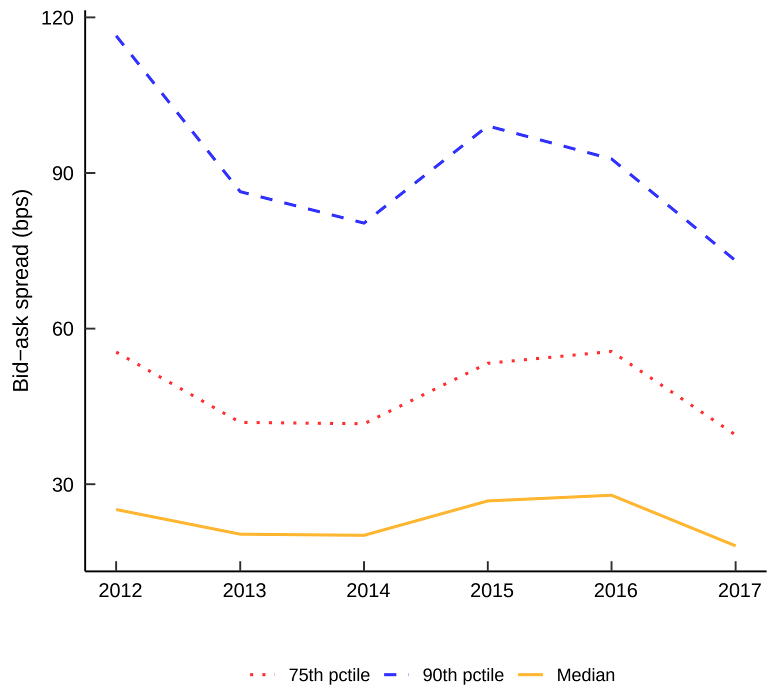


Figure 5: Dealer spreads through time

Note: This figure shows the distribution across instruments over time of bid-ask spreads earned by dealers in trades with customers. For all instruments we take all cases where the dealer sector both bought the instrument from and sold the instrument to the customer sector in the same week, and take the difference between the purchase price and the selling price. We then average this for each instrument-year pair, and plot quantiles of the distribution across instruments in each year.

Fact 5: The dealer/customer distinction embedded in models of search does not apply in the sterling corporate bond market

Much of the literature treats dealers and customers as distinct entities, with the former supplying liquidity and the latter demanding liquidity. Intermediation is thus exogenous in these models.[14] This is formalised in models by (a) assuming customers receive shocks to their asset valuations that give them a motive to trade, whilst dealers do not (Grossman

[14]Theoretical papers such as Üslü (2019), Farboodi, Jarosch and Shimer (2021) and Chang and Zhang (2020) study endogenous intermediation in theory, but papers that take search models to data typically treat intermediation as exogenous, partly due to poor data on the trades of non-dealers (Hugonnier, Lester and Weill, 2020; Brancaccio, Li and Schürhoff, 2020).

and Miller, 1988; Duffie, Gârleanu and Pedersen, 2005), and (b) assuming customers do not trade with each other. Neither of these assumptions holds in the sterling corporate bond market.

The entities that are typically identified as dealers are generally parts of universal banking groups.[15] These firms do seek to supply liquidity and make money by doing so, but have a number of other motives for trading that they share with other types of financial firm. For example, they are subject to liquidity shocks.[16] And on the other hand, firms typically treated as customers can and do supply liquidity. As described in Choi and Huh (2018) and BlackRock (2015), firms such as asset managers increasingly seek to make money by supplying liquidity in these markets. This is consistent with evidence in other markets of non-dealers supplying liquidity (Biais, Declerck and Moinas, 2016; Franzoni, Plazzi and Çötelioğlu, 2019).

A substantial amount of trade is between firms that are not typically thought of as dealers. Table 2 summarises the percentage of trades between dealers and all other traders, who we call customers. A quarter of trades do not involve a dealer as a counterparty.

As a result of these facts, our model will treat intermediation as endogenous. In particular, all traders will face shocks that lead them to demand liquidity and will be able to supply liquidity. And rather than separating traders out into dealers and customers, we will instead capture the differences between traders with heterogeneous model parameters.

Table 2: Trading by type (%)

Buyer\ Seller	Customer	Dealer
Customer	26	29
Dealer	32	13

Note: This table summarises the trading frequencies between dealers and all other traders, who we label customers.

[15] See Duffie (2010*a*) for a list of major dealer banks.

[16] Most notably, these banks are big players in derivative markets. As their derivative positions change in value banks are subject to margin calls, which mean they must post cash collateral to their counterparties. Additionally, banking regulation means bonds are both subject to capital requirements, and contribute towards banks meeting their liquidity requirements. As market conditions change – for example when a bond is downgraded – a bond's weight in these regulations changes. Both these examples give banks reason to demand, rather than supply, liquidity in bond markets.

4 A Model of Intermediation

In this section we set out a model of search and trading in a decentralised asset market. To meet our requirements and match the data, the model must have the following features: (a) traders' trading frequency must be endogenuosly determined and heterogeneous, (b) which traders intermediate must be endogenous, (c) traders' asset holdings must be unlimited and endogenous, and (d) the model must be sufficiently parsimonious to be estimable. The set of search-and-matching models building on Duffie, Gârleanu and Pedersen (2005), and in particular Üslü (2019), provides a convenient starting point. Below we set out such a model, which takes the model in Üslü (2019) and endogenises trader search intensity.

In our model, a set of traders trade a single asset. Traders face random shocks to their marginal utility of holding the asset that lead them to want to trade. Search frictions mean it takes time for a trader to find a counterparty with whom to trade. Traders can choose the rate at which they meet counterparties, but increasing this rate is costly. The costs of searching vary across traders, and at each point in time traders choose how hard to search to balance the costs of searching against the prospective gains from trade. Once two traders meet, they bargain over the trading price and quantity. An equilibrium is a steady state of this system, where search, the terms of trade, and asset holdings are the results of traders' optimising decisions.

4.1 The Setting

Agents and assets Time is continuous and the horizon infinite. There is a continuum of inifinitely-lived traders with measure 1, who discount the future at rate $r > 0$. These traders derive utility from holding a numéraire good with marginal utility equal to 1, and can hold an unconstrained amount of a durable asset with fixed supply s. A trader whose asset holding is h and net consumption of the numéraire good is c receives utility flow $u(\beta, h) + c$, where

$$u(\beta, h) = \beta h - -\kappa \frac{h^2}{2} \tag{1}$$

is the utility flow from holding the asset. A trader's lifetime utility is the present discounted value of expected utility flows, net of payments for transactions.

A trader draws a new valuation parameter β from the distribution $G(\beta)$ at Poisson arrival rate η. These random shocks create the gains from trade. We interpret them as liquidity shocks. For example, a trader who receives a negative shock could be an asset manager facing investor outflows, and thus needs to sell the asset to raise funds. A positive shock would then represent an asset manager facing investor inflows.

This utility function over holdings can be intepreted as a reduced form of a more fundamental setting where traders exhibit constant absolute risk aversion, and utility is derived from consumption rather than asset holdings (Duffie, Gârleanu and Pedersen, 2007; Üslü, 2019). In this setting, risk averse traders seeking to maximise their utility choose how much wealth to allocate between a risky asset and a risk-free asset, and how much to consume. They have a stochastic income stream which is correlated with the returns to the risky asset, and this correlation changes stochastically, changing the desirability of holding the risky asset. These shocks underly the shocks to β in $u(\beta, h)$. The parameter κ in equation (1) is a linear function of the trader's coefficient of absolute risk aversion. In the rest of this paper we will refer to κ as risk aversion. We work with the reduced-form utility, rather than the more fundamental setting, as this captures the main dynamics at play, and when we turn to estimation the parameters we can identify are those of the reduced-form utility.

Endogenous search, matching and trade Traders search for and meet counterparties with whom to trade. Traders choose the rate at which they meet counterparties, but meeting counterparties frequently is costly. Let γ denote a trader's search intensity. A trader searching with intensity γ incurs cost $s(z, \gamma)$, where parameter $z \sim F(z)$ capture a trader's cost of searching, which varies across traders and is fixed through time. The search cost function $s(z, \gamma)$ is a twice continuously differentiable, increasing and convex function of γ. Optimal search γ will be a function of the search parameter z, as well as the trader's valuation β and their current holding h.

The benefit to searching is that it results in meetings with other traders, and potential gains from trade. Meetings between traders searching with intensity γ and γ' are governed by the matching function $m(\gamma, \gamma')$, which is symmetric and linearly increasing in both arguments. This captures the fact that traders can both contact and be contacted by other traders, and that a trader can increase its meeting rate by increasing its search intensity γ.

Together, the search cost function and matching function capture the trade-off a trader faces in choosing how hard to search. Given liquidity is scarce, a trader increasing the frequency at which it can trade is beneficial, as it enables it to better optimise its portfolio and respond to shocks. But doing so, for example by hiring more people to the firm's trading desk or paying them to work overtime, is costly. Firms will choose search intensity to balance these costs and benefits. The benefits to trading are state-dependent – for example a trader may be particularly eager to trade after receiving a shock to their asset valuation, whilst after having traded they may have little desire to trade again. As a result, traders will condition their search intensity on their state – their asset holdings h and their valuation β – as well as their search parameter z.

Once two traders have been matched they engage in bilateral Nash bargaining over the quantity traded q and the per-unit price p. All traders have bargaining strength of one half. When a trader of type (z, β, h) meets a trader of type (z', β', h'), the former sells the latter $q((z, \beta, h), (z', \beta', h'))$ units of the asset at price $p((z, \beta, h), (z', \beta', h'))$.

4.2 Solving the Model

In this section we define and characterise a stationary equilibrium. Doing so requires that we analyse the three-way feedback between traders' search decisions, the distribution of the asset across traders, and the terms at which they trade. A trader's search decision depends on the distribution of the asset across traders and the terms of trade as these determine the gains to searching. Distributions depend on traders' search decisions and the terms of trade as these determine the flows of the asset between traders. Finally, the terms of trade depend on distributions and search as they determine future trading opportunities. To proceed we write down expressions for traders' value functions, their trading decisions, their search decisions and the distributions of assets across traders. An equilibrium is then a mutually consistent, stationary relationship between these endogenous variables.

4.2.1 Terms of Trade and Value Function

We begin by setting out the equations that govern the quantity and price at which traders trade, before writing down their value function. Let $V(z, \beta, h)$ be the value function of a trader with search parameter z, valuation β and holding h. We assume $V(z, \beta, h)$ is differentiable, increasing and concave. When a trader of type $\Delta = (z, \beta, h)$ meets a trader of type $\Delta' = (z', \beta', h')$, the total surplus from trading quantity q is given by the sum of the two traders' changes in value after the trade:

$$V(z, \beta, h - q) - V(z, \beta, h) + V(z', \beta', h' + q) - V(z', \beta', h') \tag{2}$$

Nash bargaining implies that the quantity they trade $q(\Delta, \Delta')$ and the price $p(\Delta, \Delta')$ solve:

$$\begin{aligned} \underset{p,q}{\text{maximise}} \quad & (V(z,\beta,h-q) - V(z,\beta,h) + pq)(V(z',\beta',h'+q) - V(z',\beta',h') - pq) \\ \text{subject to} \quad & V(z,\beta,h-q) - V(z,\beta,h) + pq \geq 0, \\ & V(z',\beta',h'+q) - V(z',\beta',h') - pq \geq 0. \end{aligned} \tag{3}$$

In words, the terms of trade maximise the product of the traders' surpluses, subject to each trader weakly prefering the agreed trade to not trading at all.[17]

The quantity that solves this problem must by definition maximise total trading surplus. The asset is thus sold by the trader who values the asset less to the trader who values it more, and the quantity sold is such that the post-trade marginal utility of the traders is equalised:

$$V_3(z,\beta,h - q(\Delta,\Delta')) = V_3(z',\beta',h' + q(\Delta,\Delta'))$$

where $V_3(z, \beta, h)$ is the derivative of the value function with respect to holdings h. Intuitively, this means trading quantity is larger when the pre-trade marginal utilities of traders are more spread out, and when the curvature of the value function is lower (as then the trading quantity that equalises traders' marginal utilities must be greater). When we estimate the model in Section 5, these facts will mean the parameters governing the slope and curvature of the value function (value β and risk aversion κ) are identified by trading quantities.

The trading surplus when two traders meet is then simply the surplus (equation (2)) evaluated at the optimal trading quantity:

$$S\big((z,\beta,h),(z',\beta',h')\big) = V(z,\beta,h - q(\Delta,\Delta')) - V(z,\beta,h) + V(z',\beta',h' + q(\Delta,\Delta')) - V(z',\beta',h')$$

The price can then be shown to be:

$$p(\Delta,\Delta') = \frac{1}{2}\left(\frac{V(z',\beta',h'+q(\Delta,\Delta')) - V(z',\beta',h')}{q(\Delta,\Delta')} + \frac{V(z,\beta,h) - V(z,\beta,h-q(\Delta,\Delta'))}{q(\Delta,\Delta')}\right)$$

The price thus depends on the slopes of the traders' value functions. As the optimal trading

[17]Given the traders' bargaining strengths are equal, they do not appear in the optimisation.

quantity approaches zero, the price approaches the average of the two traders' marginal valuations. Trading prices will therefore be more variable when traders' marginal utilities are more spread out, and when the curvature of the value function is greater (as this increases the variation in the slope of the value function). These facts will mean that trading prices, as well as trading quantities, will help to identify the value shocks β and risk aversion κ in Section 5.

We can now write down an expression for traders' value functions, taking as given their trading decisions set out above, as well as their search decisions and the distribution of the asset across traders. The Hamilton-Jacobi-Bellman equation that governs traders' optimal behaviour can be written as:

$$\begin{aligned}\underbrace{rV(z,\beta,h)}_{\text{Value}} = \underbrace{u(\beta,h) - s(z,\gamma(z,\beta,h))}_{\substack{\text{Flow value \&}\\ \text{search costs}}} + \underbrace{\eta \int (V(z,\beta',h) - V(z,\beta,h))g(\beta')d\beta'}_{\text{Switch type}} + \\ \frac{1}{2}\iiint \underbrace{m(\gamma(z,\beta,h),\gamma(z',\beta',h'))}_{\substack{\text{Meeting}\\ \text{probability}}} \underbrace{S((z,\beta,h),(z',\beta',h'))}_{\text{Surplus}} \Phi(dz',d\beta',dh') \quad (4)\end{aligned}$$

where $\Phi(z,\beta,h)$ is the cdf of traders of type (z,β,h), $\gamma(z,\beta,h)$ is the optimal search intensity for trader type (z,β,h) and $s(z,\beta,h) \equiv s(z,\gamma(z,\beta,h))$ are the corresponding search costs.

The value function is straightforward to interpret. A trader of type (z,β,h) gets flow utility $u(\beta,h)$ from holding the asset and incurs search costs $s(z,\gamma(z,\beta,h))$. At intensity η the trader draws a new value β' from $G(\beta')$. At intensity $m(\gamma(z,\beta,h),\gamma(z',\beta',h'))\Phi(dz',d\beta',dh')$ the trader meets a counterparty with type (z',β',h'), in which case they trade and extract half the total surplus.

Traders' values thus depend on the search and trading decisions of both themselves and their potential future counterparties, and the distributions of the asset across traders. A trader's search determines the search cost it incurs as well as the expected gains to trade it enjoys. The distribution of the asset across traders, as well as these traders' search decisions, then determine how frequently the trader meets a counterparty, and which type of counterparty they meet. The Nash bargaining protocol then determines the gains to trade when they do meet a counterparty.

4.2.2 Trader Search

We now characterise traders' optimal search decisions, given their trading decisions and the distribution of the asset. To do so, we take the first-order condition of the value function in equation (4) and apply the envelope theorem, yielding the equation governing the optimal search of type $\Delta = (z, \beta, h)$:

$$\underbrace{s_2(z, \gamma(\Delta))}_{\substack{\text{Marginal cost} \\ \text{of search}}} = \frac{1}{2} \int \underbrace{\frac{\partial m(\gamma(\Delta), \gamma(\Delta'))}{\partial \gamma(\Delta)}}_{\substack{\text{Increase} \\ \text{in meetings}}} \underbrace{S(\Delta, \Delta')}_{\substack{\text{Surplus from} \\ \text{meeting}}} \Phi(d\Delta') \tag{5}$$

The left-hand side is the marginal cost of searching, which is increasing in the trader's search intensity. The right-hand side is the marginal benefit to searching. Increasing search intensity increases the rate at which the trader meets a counterparty and enjoys the expected gains from trade. The expected gains from trade depend on the trading surplus when the trader meets a given counterparty, and the likelihood of meeting each counterparty type. Traders thus choose their search intensity to balance the gains from trade and the costs of search, and will search harder if their search costs are lower, or if the expected gains from trade are higher.

4.2.3 Type Distributions

We now close the model by providing expressions for the equilibrium distributions of assets across trader types. For the system to be in steady state, the net flows from trading and value shocks must be equal to zero across all trader types (z, β, h). The measure of traders with search costs z, valuation $\beta \leq \beta^*$ and holding $h \leq h^*$ must satisfy:

$$\begin{aligned} &\iint_{\underline{\beta}}^{\beta^*} \int_{h^*}^{\infty} m(\gamma(z,\beta,h), \gamma(\Delta'))\phi(z,\beta,h)\phi(\Delta')\mathbb{1}(q(z,\beta,h,\Delta') \geq h - h^*) dh d\beta d\Delta' - \\ &\qquad \iint_{\underline{\beta}}^{\beta^*} \int_{-\infty}^{h^*} m(\gamma(z,\beta,h), \gamma(\Delta'))\phi(z,\beta,h)\phi(\Delta')\mathbb{1}(q(z,\beta,h,\Delta') < h - h^*) dh d\beta d\Delta' \\ &\qquad\qquad = \eta(1 - G(\beta^*)) \int_{\underline{\beta}}^{\beta^*} \int_{-\infty}^{h^*} \phi(z,h,\beta) dh d\beta - \eta G(\beta^*) \int_{\beta^*}^{\bar{\beta}} \int_{-\infty}^{h^*} \phi(z,h,\beta) dh d\beta \end{aligned} \tag{6}$$

for all (z, β^*, h^*). The left-hand side represents net inflows due to trade. This consists of the flow of traders with type $(z, \beta \leq \beta^*, h > h^*)$ who meet a counterparty and sell enough of the asset such that their holdings fall below h^*, minus the flow of traders with type $(z, \beta \leq \beta^*, h \leq h^*)$ who meet a counterparty and buy enough that their holdings exceed h^*. The right-hand side consists of net outflows due to value shocks. The outflows consist of the stock of traders with type $(z, \beta \leq \beta^*, h \leq h^*)$, who with intensity η receive a new value shock, and with probability $1 - G(\beta^*)$ draw a new value above β^*. The inflows are traders with type $(z, \beta > \beta^*, h \leq h^*)$ who with intensity $\eta G(\beta^*)$ receive a shock that takes their value below β^*.

Finally, the distribution of types in equilibrium must be consistent with the ex ante distribution of types:

$$\iint \Phi(z, \beta, h) d\beta dh = F(z) \quad \forall z \tag{7}$$

4.2.4 Equilibrium

We now define an equilibrium based on the system of equations derived above. An equilibrium is a set of value functions, search intensities, terms of trade and distributions of agents and assets that are mutually consistent in steady state. Let the space of trader types (z, β, h) be $\mathcal{T} = \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}$. A steady state equilibrium is:

- value function $V : \mathcal{T} \to \mathbb{R}$;
- pricing function $p : \mathcal{T}^2 \to \mathbb{R}^+$;
- trading quantity function $q : \mathcal{T}^2 \to \mathbb{R}$;
- density function $\phi : \mathcal{T} \to \mathbb{R}^+$; and
- search intensity functions $\gamma : \mathcal{T} \to \mathbb{R}^+$

that solve the following equilibrium conditions:

- value functions (equation (4));
- distribution functions (equation (6) and (7));
- optimal search (equation (5));
- optimal prices and quantity (equation (3)); and
- market clearing:

$$\iiint h\Phi(dz, d\beta, dh) = s$$

An equilibrium is thus a set of equations that solves the three-way feedback between traders' search decisions, the distribution of the asset across traders, and the terms at which they trade. We solve for this equilibrium numerically in the sections that follow. If we fix each trader's search intensity exogenously, our model becomes that of Üslü (2019). He shows that in this case an equilibrium exists and is unique. Based on our numerical solutions in the rest of the paper, the addition of endogenous search does not change this.

4.3 Properties of the Equilibrium

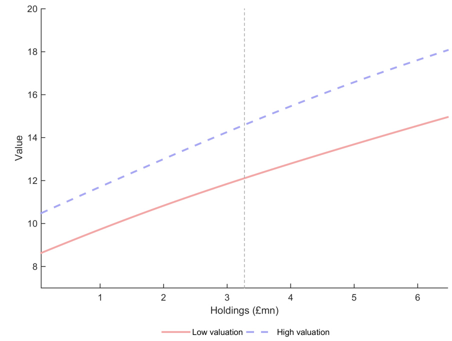
Figure 6 summarises the equilibrium, showing the value functions, prices, trading quantity and distributions of traders with different valuations β, for the parameter vector we estimate in Section 5. The value function is increasing and concave, inheriting these properties from the quadratic flow utility (Figure 6a).[18] A trader with a low valuation β values the asset less, and so at any level of holding is more likely to be a net seller than a trader with a high valuation (Figure 6b), and trades the asset at a lower price (Figure 6c). As a result of these differences in trading patterns, the low-value trader holds less of the asset (Figure 6d). With quadratic utility and a symmetric distribution of β, the equilibrium distributions are symmetric around the per-capita supply of the asset s, with the distributions of holdings for traders with low valuation the mirror image of those with high valuation.

4.3.1 Search and Trading

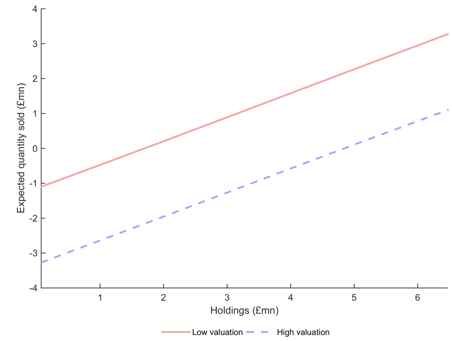
Figure 7a summarises how traders optimally search, showing how a given trader varies their search and trading behaviour depending on their valuation and asset holdings. For a given valuation β a trader's search intensity is a convex function of its asset holdings. Search intensity takes its minimum at a quantity which we call the trader's *target holding*, defined as the level of asset holding $h^*(z, \beta)$ at which the returns to search are minimised.

Trader search thus takes an intuitive form: traders search least when the gains from trade are low, and search harder as the gains from trade increase. Concave utility naturally means that in equilibrium traders have more to gain from a trade when they have very low or high holdings. Search intensity is thus U-shaped in holdings. This is the mechanism that lies behind Figure 2, which shows that traders offset their trades far more frequently

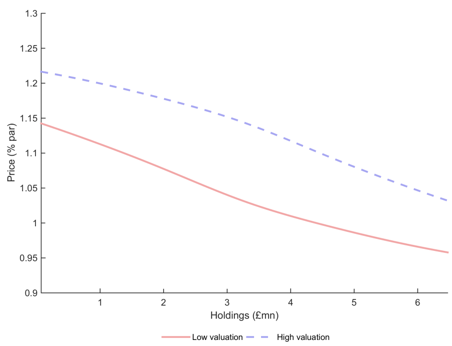
[18] This concavity is necessary for the model to have a solution, as it ensures traders will always trade and hold a finite amount of the asset.



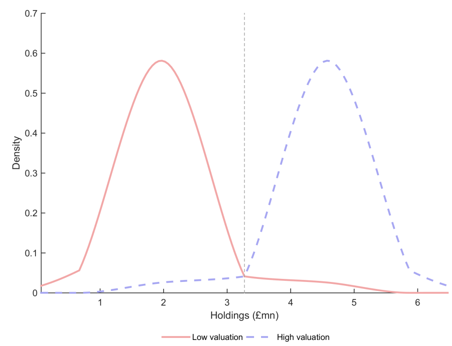
(a) Value functions



(b) Average quantity sold



(c) Average trading price



(d) Distributions

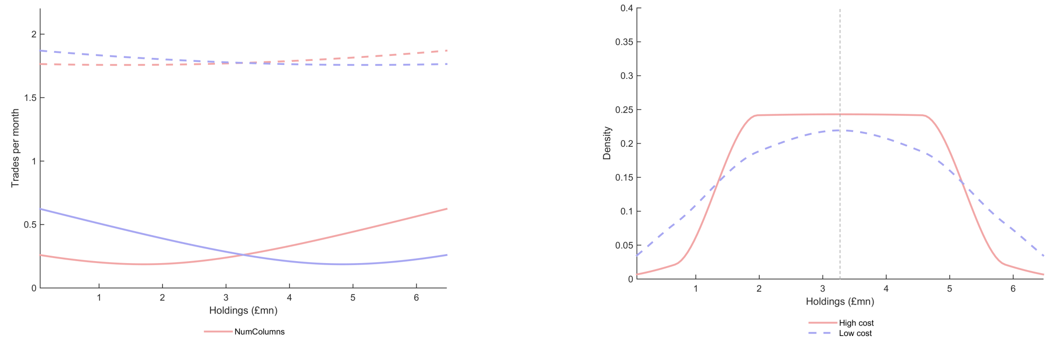
Figure 6: Terms of trade and distributions by trader type

Note: These figures summarise key features of the model's equilibrium, for the parameter values estimated in Section 5. Each panel plots one of the model's outcome variables against a trader's holdings for different levels of value shock β, for a given search parameter z. Panel (a) shows the value function $V(z, \beta, h)$. Panel (b) shows the average quantity sold per trade. Panel (c) shows the average trading price. Panel (d) shows the distribution of holdings for type (z, β). The grey vertical lines show per-capita supply of the asset.

than if their search was exogenous. When search is chosen optimally, traders increase search intensity whenever a trade or a shock takes them away from their target holdings.

Traders with low search costs (or high search efficiency z) search harder, and trade more frequently as a result (Figure 7a). Their search intensity is less sensitive to their value β. Put differently, the difference in search intensity between low- and high-cost searchers is greatest when traders are at their target holdings, and have no reason to demand liquidity. Thus low-cost traders do not leverage their cost advantage to respond faster to shocks, but instead to supply liquidity.

Figure 7b shows how a trader's search technology influences its asset holdings. Traders



(a) Trading frequency by shock & holding

(b) Holding distribution by type

Figure 7: Search and distributions by trader type

Note: These figures summarise some key quantities in our estimated model. Panel (a) shows trading frequency by asset holding h and shock value β, for a trader with a given search efficiency z. Panel (b) shows the distribution of agents across different holdings for low-search-cost traders (high z) and high-search-cost traders (low z). The grey vertical lines show per-capita supply of the asset.

who are able to search more efficiently allow their holdings distribution to spread out more.[19] They are willing to trade to more extreme holdings because their superior search technology means they will be able to offset this trade relatively quickly.

4.3.2 Liquidity

Liquidity in this context is the ease with which traders can change their holdings when they wish to. Traders who wish to change their asset holdings demand liquidity, and those who facilitate this supply liquidity. Traders demanding liquidity face the following types of illiquidity: (a) the time they must wait to trade, (b) the cost they pay to search, (c) the extent to which the quantity they trade is rationed relative to what they would trade in a market without frictions, and (d) the difference in the price at which they trade relative to that in a market without frictions. Aggregate welfare is determined by how the assets are distributed across traders with different values and the search costs traders incur, with trading profits netting out across traders.

The model gives us a flexible framework with which to study search and liquidity in over-the-counter markets. It gives predictions about the key variables of interest in trading markets – prices, quantities, holdings and trading frequencies – and correspondingly gives traders a relatively rich set of variables they can adjust in response to shocks and in counter-

[19] Üslü (2019) finds a similar result in a setting with exogenous trading frequency.

factual scenarios. The fact that holdings and trading quantity are unrestricted and search is endogenous means we can study both the extensive and intensive margins of trading, namely whether a trade takes place and how much is traded, and how traders use these margins to manage their asset portfolios.

The key outcome variables of the model depend on the values of its parameters. For example average prices depend on the slope of the value functions, the variance of prices depends on how much these slopes vary, and the trading frequency depends on the cost of searching. Similarly, welfare, liquidity and the impact of counterfactual scenarios will depend on the model's parameters. To answer our key questions, then, we need to estimate the model.

5 Estimation

We first set out the parametric assumptions we make to take the model to the data. We then set out our estimation procedure, before summarising the moments we use. Finally we set out the key variation that identifies our models' parameters.

5.1 Parametric Assumptions

We use the following linear matching function:

$$m(\gamma, \gamma') = 2\gamma \frac{\gamma'}{\Gamma} \tag{8}$$

where

$$\Gamma = \iiint \gamma(z, \beta, h) \Phi(dz, d\beta, dh)$$

This technology captures the fact that a trader can both contact and be contacted by another trader. Conditional on contact the likelihood of meeting a given type of counterparty is proportional to their search intensity.[20]

[20] Another common linear matching function is $m(\gamma, \gamma') = \gamma + \gamma'$ (Shimer and Smith, 2001), which assumes that counterparties are drawn uniformly from the distribution of traders. We do not use this functional form because (a) trading frequency is heavily right-skewed as shown in Figure 1 and (b) many traders have a monthly trading frequency per instrument that is close to 0. As a result, this matching function would imply

We take the search cost function to be:

$$s(z, \gamma) = (\gamma - z)^2 \tag{9}$$

where $z \sim F(z)$. A trader can costlessly meet traders at rate z, but must pay a constant marginal cost to meet traders more frequently. Intuitively, z can be thought of as a flow of contacts a trader makes as part of its everyday business. For example, if the trader is an investment bank it meets its clients regularly as part of its broader investment banking activity. However, to increase its meeting rate above this minimum level is costly, as it must contact other traders or hire more staff to make the contacts. For simplicity, we take the marginal cost of these contacts to be linear.

We assume the search parameter z follows a Gamma distribution, with shape parameter k_z and scale parameter θ_z. We choose this distribution as it can match the skewness of the trading rate distribution (Figure 1), has strictly positive support, and is relatively parsimonious. We assume the shock distribution $G(\beta)$ is uniform, with mean μ_β and variance σ^2_β.

5.2 Estimation Procedure

We estimate the model using the data described in section 2.1, assuming these data are generated by the model in steady state. The unit of time is 1 month and the monthly discount rate is 0.5%. We fix s to be equal to the mean total amount outstanding in a bond, normalised by the number of traders trading in that bond.

We estimate the parameter vector $\psi = \{\mu_\beta, \sigma_\beta, \kappa, \eta, k_z, \theta_z\}$, where $\{\mu_\beta, \sigma_\beta\}$ are the parameters of the β distribution, κ is risk aversion, η is shock frequency and $\{k_z, \theta_z\}$ are parameters of the distribution of the search parameter z. We use a minimum-distance estimator that matches theoretical moments implied by the model to their empirical counterparts. In practice this takes the form of a nested loop: for any given ψ we numerically solve the model. We then calculate a vector of theoretical moments $m(\psi)$ and compare this to its empirical counterpart m_0. We choose the parameter vector that minimises the sum of squared differences between the theoretical and empirical moments:

negative search intensities for the left-tail of the trading frequency distribution.

$$\hat{\psi} = \arg\min_{\psi} (m(\psi) - m_0)'(m(\psi) - m_0)$$

To find the model's equilibrium we solve the model's equations at a finite set of points in the type space $\mathcal{T}$. The distributions of valuations β and search cost parameters z are discrete versions of those set out in section 5.1. We use interpolation splines to ensure continuity of functions in holdings h. We then use a nonlinear solver to numerically solve the equilibrium equations. For full details of the estimation procedure, see Appendix A2.

5.3 Moments

Our data contain the key information we need to identify the model: trading frequencies, the terms of trade and asset holdings. The data contain the identity of all traders, and not just dealers as is commonly the case, which means we can characterise trader heterogeneity across all traders. In this section we set out the moments we use to estimate the model's parameters.

We match two sets of moments: *across-trader* moments and *within-trader* moments. Within-trader moments measure the distribution of quantities for a given trader – for example how much its inventory varies through time. Across-trader moments measure the distribution of average quantities across traders – for example comparing how the average trading frequency of one trader differs from the average trading frequency of other traders.[21] We choose moments covering the means, variances and correlations of traders' trading frequencies, trading prices, trading quantities and holdings to identify our parameters. We give analytical expressions for each of these moments below.

5.3.1 Theoretical Moments

Expectations

1. Average trading frequency.

$$\mathbb{E}(n) = 2 \int \gamma(\Delta)\phi(\Delta)\pi(\Delta)d\Delta$$

[21]Note that the means of a quantity are the same within and across traders.

where $\pi(\Delta)$ is the fraction of Δ's meetings that result in a trade:

$$\pi(\Delta) = \int \mathbb{1}(q(\Delta, \Delta') \neq 0) \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta'$$

2. Average trade size.

$$\mathbb{E}(|q|) = \int \frac{\gamma(\Delta)\phi(\Delta)}{\Gamma} \int |q(\Delta, \Delta')| \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\Delta$$

3. Average price.

$$\mathbb{E}(p) = \int \frac{\gamma(\Delta)\phi(\Delta)}{\Gamma} \int p(\Delta, \Delta') \frac{\mathbb{1}(q(\Delta, \Delta') \neq 0)}{\pi(\Delta)} \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\Delta$$

Across-trader moments

4. Standard deviation of trading frequency across traders.

$$SD^A(n) = \sqrt{\int (n(z) - \mathbb{E}(n))^2 f(z) dz}$$

where $n(z)$ is type z's average trading frequency:

$$n(z) = 2 \iint \gamma(z, \beta, h) \frac{\phi(z, \beta, h)}{f(z)} \pi(z, \beta, h) d\beta dh$$

Within-trader moments

5. Standard deviation of holdings, within traders.

$$SD^W(h) = \int \sqrt{\mathbb{V}(h|z)} f(z) dz$$

where $\mathbb{V}(h|z)$ is the variance in holdings for a trader with search parameter z:[22]

$$\mathbb{V}(h|z) = \iint (h - s)^2 \frac{\phi(z, \beta, h)}{f(z)} d\beta dh$$

[22]Note that the symmetry of the equilibrium shown in Figure 6d means that each trader on average holds the asset's per-capita supply, s.

6. Standard deviation of trading prices, within traders.

$$SD^W(p) = \int \sqrt{\mathbb{V}(p|z)} f(z) dz$$

where $\mathbb{V}(p|z)$ is the variance of trading prices for a trader with search parameter z:

$$\mathbb{V}(p|z) = \iint \frac{\gamma(z,\beta,h)\phi(z,\beta,h)}{\Gamma(z)f(z)} \times \\ \int \big(p(z,\beta,h,\Delta') - \mathbb{E}(p|z)\big)^2 \frac{\mathbb{1}(q(z,\beta,h,\Delta') \neq 0)}{\pi(z,\beta,h)} \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\beta dh$$

$\mathbb{E}(p|z)$ is type z's mean trading price:

$$\mathbb{E}(p|z) = \iint \frac{\gamma(z,\beta,h)\phi(z,\beta,h)}{\Gamma(z)f(z)} \times \\ \int p(z,\beta,h,\Delta') \frac{\mathbb{1}(q(z,\beta,h,\Delta') \neq 0)}{\pi(z,\beta,h)} \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\beta dh$$

and $\Gamma(z)$ is type z's average search intensity:

$$\Gamma(z) = \iint \gamma(z,\beta,h)\phi(z,\beta,h) d\beta dh$$

7. Correlation between quantity sold and holdings, within-traders.

$$corr^W(h,q) = \int \frac{cov(h,q|z)}{\sqrt{\mathbb{V}(h|z)\mathbb{V}(q|z)}} f(z) dz$$

where expressions for the conditional (trade-weighted) covariance of holdings and quantity sold $cov(h,q|z)$, holdings variance $\mathbb{V}_{TW}(h|z)$ and variance of quantity sold $\mathbb{V}(q|z)$ are given in Appendix A3.

8. Correlation between absolute inventory $inv \equiv |h - s|$ and trading frequency, within traders.

$$corr^W(inv,n) = \int \frac{cov(inv,n|z)}{\sqrt{\mathbb{V}(inv|z)\mathbb{V}(n|z)}} f(z) dz$$

where expressions for the conditional covariance of absolute inventory and trading frequency $cov(inv, n|z)$, variance of absolute inventory $\mathbb{V}(inv|z)$ and variance of trading frequency $\mathbb{V}(n|z)$ are given in Appendix A3.

5.3.2 Empirical Moments

We compute the empirical counterpart of each of the moments set out above to construct the empirical moment vector m_0. The model is defined at the instrument level, so each of the moments is computed for a given instrument. We then average the moments across instruments to produce the vector m_0. The size of our sample is then the number of instruments in our data. To compute bootstrap standard errors, we resample from the set of instruments and repeat the estimation for each bootstrap sample.

The data are likely to include the effects of economic forces outside of our model. We attempt to cleanse the data of these other factors by including a number of fixed effects. In particular, our within-trader moments are equivalent to including trader-level fixed effects and computing moments based on the residual variation. This strips out any time-invariant trader-specific variation from these moments. For example if a trader holds more of a bond throughout its life as it was an undwriter on the bond's initial issuance, we strip this variation out when computing the variance of trader holdings. When computing the within-trader variance in prices, we further control for the current credit rating of the instrument, and compute the variation in prices based on the residual variance after controlling for the credit rating. We do this to control for the effect of news about the fundamentals of an instrument on its price.

5.4 Identification

The moments set out above provide a mapping from the model's parameters to the data. In this section we show that there is sufficient variation to invert this mapping and infer the parameters from the data. The model's nonlinearity means that all but one parameters affect all moments, but for each parameter we can set out the key variation that pins it down. Table 3 summarises this key variation for each parameter.

The search parameters z are straightforward to identify, as the optimal search equation (5) ensures that a trader's average trading frequency is a monotonic function of its search parameter z. As a result the parameters of the distribution of z are pinned down by the moments of the distribution of traders' average trading frequencies.

Table 3: Identification

Parameter	Moment
$\mathbb{E}(z) = k_z\theta_z$	$\overset{[+]}{\mathbb{E}(n)}$
$\mathbb{V}(z) = k_z\theta_z^2$	$\overset{[+]}{SD^A(n)}$
μ_β	$\overset{[+]}{\mathbb{E}(p)}$
σ_β	$\overset{[+]}{SD^W(p)}$, $\overset{[+]}{\mathbb{E}(\lvert q\rvert)}$, $\overset{[+]}{SD^W(h)}$
κ	$\overset{[+]}{SD^W(p)}$, $\overset{[-]}{\mathbb{E}(\lvert q\rvert)}$, $\overset{[-]}{SD^W(h)}$
η	$\overset{[+]}{corr^W(\lvert h-s\rvert, n)}$, $\overset{[+]}{corr^W(h,q)}$

Note: This table summarises the key variation that identifies our parameters. n is the number of trades, p is the trading price, q is the amount sold by a trader and h is a trader's holding. $SD^W(x)$ is the standard deviation of x for a given trader, whilst $SD^A(x)$ takes the average value of x for each trader and computes the standard deviation of this across traders. $corr^W(x, y)$ is the correlation of x and y for a given trader. Shocks β are distributed according to a uniform distribution with mean μ_β and variance σ_β^2. Parameter κ governs traders' risk aversion. Parameter η is the frequency (per month) at which traders draw new values β. Parameter z in search cost function $z = (\gamma - z)^2$ is distributed according to a Gamma distribution with shape parameter k_z and scale parameter θ_z. A [+] above a moment indicates the moment is increasing in the relevant parameter, whilst a [−] indicates it is decreasing.

The mean of the value shock distribution μ_β is similarly straightforward to identify. This is because it affects the mean price, but no other moments in the model. As set out in Section 4, when two traders meet, the quantity q they trade is such that their post-trade marginal valuations for the asset are equalised: $V_3(z, \beta, h-q) = V_3(z', \beta', h'+q)$. Moving the location of the valuation distribution has no effect on this, as it affects all traders equally. As a result, the location of the β distribution doesn't affect trading surplus, and so doesn't affect search intensity or the asset distributions. The average trading price, which depends on the average slope of traders' value functions, does depend on the location of β, and as a result pins it down exactly.[23]

The variance of the value shock σ_β^2 and the level of risk aversion κ are identified by the distributions of holdings, trade sizes and price. A decrease in σ_β^2 causes the value functions across different levels of β to move closer together, whilst an increase in risk aversion increases the curvature of the value function in holdings for a given β. Both these shifts cause traders

[23] In practical terms this makes estimation a simpler task, as we can estimate the mean of the shock distribution separately from the other parameters. In a first stage we search over $(\sigma_\beta, \kappa, \eta, k_z, \theta_z)$ to match all moments except the mean price. In a second stage we choose μ_β to perfectly fit the mean price.

to shrink the variance of their asset holdings. Intuitively, both a smaller shock variance and increased risk aversion reduce a trader's incentive to trade to extreme asset holdings. They also both decrease the trade size. As explained in Section 4 trade size is smaller when traders' marginal valuations are less spread out (as is the case when σ^2_β is small) and when the curvature of the value function is greater (as is the case when κ increases).

An increase in risk aversion κ and a decrease in the variance of shocks σ^2_β have opposite predictions, however, for the variance of prices. As set out in Section 4, the price at which two traders trade is governed by the slopes of their value functions. A reduction in σ^2_β reduces the variance of these slopes across different β and hence reduces the variance of prices. An increase in κ, however, increases the curvature of utility for a given β, which increases the variation in the slope of the value function across holdings. This increases the variance of prices, and enables us to separate the variance of shocks from traders' risk aversion.

The shock frequency, η, is identified by the correlation between traders' holdings of the asset and their trading price and quantity. An increase in the shock frequency moves the value functions of traders with different levels of β closer together. As a trader's value becomes more transient, they place less weight on their current value in their trading decisions and more on their expected future value, which is simply the mean of the distribution of β. In the context of Figure 6b and Figure 7a, this causes the average trading quantities and trading frequencies of traders with different β to move closer together. This strengthens the positive correlation between their holdings and the amount they sell, and their inventories – defined as the absolute difference between their holdings and the per-capita supply of the asset – and the amount they trade. These moments thus pin down the frequency of shocks.

6 Results

Table 4 shows the estimated parameters. The search parameter z is highly heterogeneous and highly skewed (Figure 8). The bulk of traders have low values of z, meaning they find it costly to meet people often. A minority of traders have high values of z, which means they have a low cost of searching and are able to meet counterparties frequently. This difference in search technology is what lies behind the heterogeneous trading rates seen in Figure 1, and the fact that over-the-counter markets often seem to have a core of connected, active traders and a periphery of less connected traders (Di Maggio, Kermani and Song, 2017; Mallaburn, Roberts-Sklar and Silvestri, 2019).[24]

Table 4: Parameter Estimates

Parameter	Estimate
Search efficiency $z \sim \Gamma(k_z, \theta_z)$	
k_z	0.545
θ_z	0.374
Shock frequency η	
η	0.040
Utility $u(h) = \beta h - 0.5\kappa h^2$; $\beta \sim U(\mu_\beta, \sigma_\beta)$	
μ_β	0.031
σ_β	0.015
κ	0.008

Note: Parameter z in search cost function $z = (\gamma - z)^2$ is distributed according to a Gamma distribution with shape parameter k_z and scale parameter θ_z. Parameter η is the frequency (per month) at which traders draw new values β. Shocks β are distributed according to a uniform distribution with mean μ_β and variance σ_β^2. Parameter κ governs traders' risk aversion.

Shocks are infrequent – a trader faces a shock on average 0.5 times a year. These infrequent shocks are consistent with the infrequent trading we observe in corporate bond markets. Traders trade more frequently – 0.4 times a month (Figure 5) – than they are shocked. In a frictionless market these two frequencies would be the same. This difference can be explained by two factors: (a) when it meets a counterparty, a trader is unable to trade the full amount it would like, and will thus trade again, and (b) traders don't just trade to demand liquidity when shocked, but also to supply liquidity to other traders.

The parameters of the utility function $u(\beta, h) = \beta h - 0.5\kappa h^2$ are reduced-form parameters, and thus their magnitudes are difficult to interpret in terms of fundamental preferences.[25] Here we limit ourselves to a discussion of the relative magnitudes of the utility parameters β and κ, and note that below we will show the model fits the data very well. Both the variance of β and the level of κ govern variation in the slope of the utility function: κ creates variation in marginal utility across holdings for a given β, whilst a shock to β changes the marginal

[24]Institutionally, these active traders will be firmslarge investment banks for example – that have large client networks who they contact regularly. This large client network could come from some more efficient technology – these traders are simply more effective at being middlemenor could be due to the fact that they often have other business lines with clients – for example derivatives tradingand can leverage these contacts in bond trading. The less active traders don't enjoy the same technologies or client contacts, and are thus less able to trade frequently.

[25]As explained in Section 4 $u(\beta, h)$ is a reduced-form version of a more fundamental model with CARA preferences over consumption. However, given we cannot identify the parameters of this more fundamental setup, it is hard to map our parameter estimates into this setting in a quantitative sense.

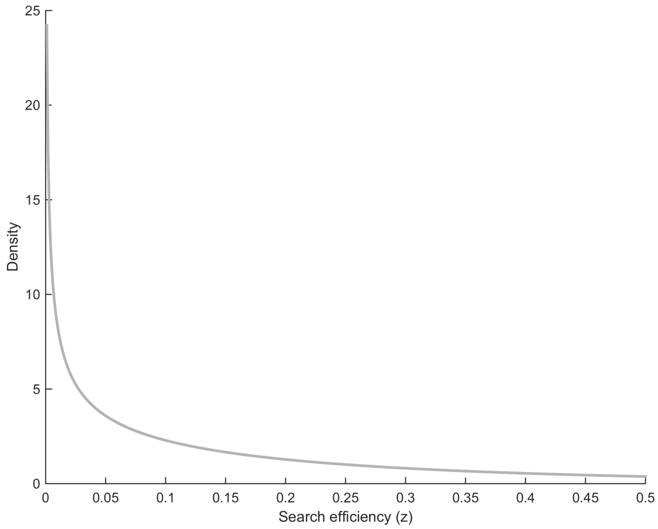


Figure 8: Distribution of search parameter z

Note: This chart shows the estimated distribution of the parameter z in traders' search cost function $s(z, \gamma) = (\gamma - z)^2$, where z follows a Gamma distribution. A trader with a high z has low search costs.

utility for a given holding. The variation caused by β is quantitatively larger than that caused by κ. Concretely, a one standard deviation shock to β causes a change in marginal utility that is 50% greater than that caused by a one standard deviation change in holdings, which operates via κ. The concavity of a trader's utility function, which is governed by κ, is thus small relative to the distance between the utility functions of traders with different valuations β. As a result, transferring the asset between traders with the same valuation has a limited effect on utility. This feature of our estimates – the limited concavity of trader utility – will be key in driving our results in the counterfactuals.

Table 5 summarises the model fit. The model fits price, holdings, trading quantity and trading frequency moments well. Two moments merit further discussion. In the data, variation in price is more limited than variation in asset holdings: the standard deviation of the prices a trader pays for a bond is around 4% of the mean, whilst the standard deviation of their holdings is 40% of the mean. Fitting these two moments is what drives the risk aversion parameter κ to be quite low in our results, as low risk aversion reduces the variability of prices but increases the variability of holdings.

Table 5: Model fit

Moment	Data	Model
Mean price, %	1.09	1.09
Std. dev. price within traders, %	0.04	0.04
Mean holdings, £mn	3.27	3.27
Std. dev. holdings within traders, £mn	1.23	1.31
Mean trade size, £mn	0.67	0.55
Mean trading frequency, per month	0.44	0.43
Std. dev. trading frequency across traders, per month	0.55	0.55
Correlation inventory & trading frequency	0.08	0.09
Correlation holdings & quantity sold	0.33	0.30

Note: This table shows empirical moments and simulated moments calculated using the parameter estimates in Table 4.

6.1 Liquidity & Intermediation

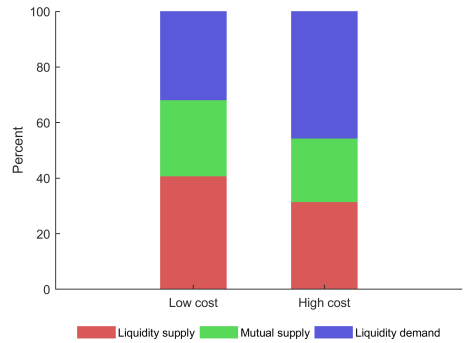
Table 6 summarises liquidity in the market according to our estimated model. On average traders are £0.45mn away from their target holding $h^*(z,\beta)$, defined as the level of holdings at which the expected gains to trade for type (z,β,h) are minimised. After a shock, traders wait an average of 2 months before they're able to trade. And once able to trade, the quantity they trade is rationed: on average they only trade 28% closer to their target holding. In a frictionless market, by contrast, traders are always at their target holding, can trade instantly after a shock, and trade to their new target holding in a single trade.

Table 6: Liquidity statistics

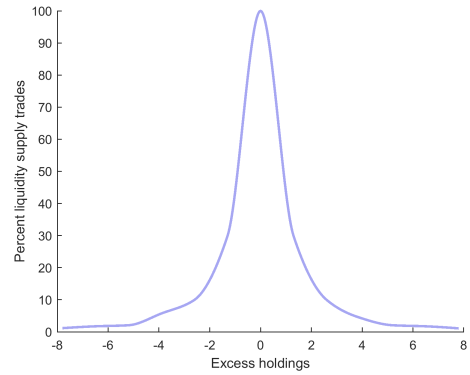
	Mean
Distance from target, £mn	0.45
Waiting time, months	2.0
Quantity rationing (%)	0.72

Note: This table summarises liquidity according to our estimated model. Distance from target is the mean of traders' excess holdings $h - h^*(z,\beta)$, where $h^*(z,\beta)$ is a trader's target holding. Waiting time is the average period before a trader meets a counterparty after a shock that changes their valuation. Quantity rationing is the average of trader's post-trade distance from target holding as a percentage of their pre-trade distance from target holding.

A trader supplies liquidity to their counterparty if they enable their counterparty to trade closer to their target holding position. Formally, a trader of type (z,β,h) supplies liquidity to their counterparty of type (z',β',h') if the following object is negative, and demands it if it is positive:



(a) Intermediation by search cost



(b) Intermediation by excess holdings

Figure 9: Intermediation by trader type and excess holdings

Note: These figures summarise intermediation in our estimated model. Panel (a) shows the fraction of trades that supply liquidity, demand liquidity and mutually supply liquidity for low-search-cost traders (high z) and high-search-cost traders (low z). Panel (b) shows how the fraction of trades that supply liquidity depends on a trader's excess holdings $h - h^*(z, \beta)$.

$$|h' + q((z, \beta, h'), (z', \beta', h')) - h^*(z', \beta')| - |h' - h^*(z', \beta')| \tag{10}$$

where $h^*(z, \beta)$ is the target holding for type (z, β, h).

Figure 9 summarises intermediation by trader type and by trader inventory. Figure 9a shows that traders with low search costs supply liquidity more than traders with high search costs. If any trader were to be described as an intermediary, it would thus be the trader with low search costs. However, intermediation is crucially a trade-specific concept: all types of traders supply liquidity in the course of their trading. Figure 9b shows how the extent to which traders supply liquidity varies with their excess holdings, defined as their holding h minus their target holding $h^*(z, \beta)$. Traders at their target holdings are more prone to supply liquidity than traders with large positive or negative inventories.

7 Counterfactuals

7.1 Regulating Dealers

Tighter capital regulation increases the fraction of a bank's balance sheet that must be funded by equity, rather than debt. If a bank applies this evenly across all its business, this increases the fraction of any position - short or long - that must be funded by equity.[26]

There are two mechanisms by which tighter capital regulation can reduce a bank's incentives to hold assets. The first is via violations of the Modigliani-Miller theorem (Modigliani and Miller, 1958). A bank's cost of equity is typically higher than its cost of debt. In the absence of any frictions, increasing a bank's equity funding will not increase its overall cost of capital, as its costs of equity and of debt will fall. In a world with frictions, its cost of capital will increase.[27] The second is via a violation of the debt overhang problem (Myers, 1977) discussed by Andersen, Duffie and Song (2019) and Duffie (2018), by which capital regulation reduces a bank's willingness to take a new, relatively safe asset onto its balance sheet.

Given dealer-affiliated banks have typically acted as intermediaries in fixed income markets, there are concerns that bank capital regulation could harm market functioning. We model tighter capital regulation as an increase in the cost of holding a long or short position in an asset for low-cost traders. We apply this cost increase to low-cost traders in our model, as these are the traders that intermediate. Their flow utility function becomes:

$$u(\beta, h) = \beta h - \kappa \frac{h^2}{2} - \tau |h| \tag{11}$$

where $\tau \geq 0$ represents the increase in costs due to regulation.

Table 7 summarises how market quantities change for $\tau = 1\%$. Tighter regulation of dealers leads them to significantly decrease their holdings of the asset, as they find holding inventory more costly. Market clearing implies unregulated traders increase their holdings. Tighter regulation, by reducing the marginal utility of holding the asset for dealers, reduces the average trading price.

Dealers adjust their trading frequency in two ways, as shown in Figure 10. Firstly, their target holding - where search is at its minimum - shifts to the left, reflecting their reduced

[26]Whilst capital regulation applies at the bank level, banks' internal capital markets are typically organised such that increases in capital requirements take effect across the bank's business lines (Bajaj, Binmore, Dasgupta and Vo, 2018).

[27]Potential violations of the Modigliani-Miller assumptions include the tax advantage of debt (Kashyap, Stein and Hanson, 2010), asymmetric information leading to a 'pecking order' theory of financing (Myers and Majluf, 1984), agency problems of bank management (Diamond and Rajan, 2001) and the possibility that banks' short-term debt enjoys a 'money-like' convenience yield (Stein, 2012; Kashyap, Stein and Hanson, 2010). Note that in the case of the tax shield on debt, a full welfare analysis would need to take into account the increase in tax revenues resulting from higher capital requirements (Admati, DeMarzo, Hellwig and Pfleiderer, 2013).

Table 7: Capital counterfactual

	% change in counterfactual
Dealers	
Expected holdings	-33
Spread received	17
Expected utility	-37
Other traders	
Expected holdings	6
Spread paid	17
Expected utility	1
Aggregate	
Price	-19
Trading frequency	1
Trade size	4
Waiting time	-5
Quantity rationing	-2
Aggregate welfare	-5

Note: This table summarises the changes in key variables in the capital counteractual, relative to our baseline equilibrium. Capital regulation is modelled as setting $\tau = 1\%$ for low-cost traders, where their flow utility from holding the asset is given by $u(\beta, h) = \beta h - 0.5\kappa h^2 - \tau|h|$. Waiting time is the average period before a trader meets a counterparty after a shock that changes their valuation. Quantity rationing is the average of trader's post-trade distance from target holding as a percentage of their pre-trade distance from target holding.

desire to hold inventory. Secondly, the slope of the trading frequency curve becomes steeper when they are away from their target, meaning regulation leads them to search harder when away from target inventory. This is the mechanism driving the empirical finding in Figure 3 that dealers' tendency to offset trades is related to capital regulation, and stems from the fact that regulation has increased the costs of being away from target holding.

Tighter regulation increases the spreads earned by dealers (Table 7). This reflects the balance of two forces. On the one hand, regulation means dealers are less inclined to trade away from target holding, as the costs to taking large positions are greater. All else equal, this would increase the spread required to persuade them to do so. On the other hand, conditional on being away from target holding, dealers are more anxious to trade back to target. This would decrease the spread they pay, as in these cases they are demanding, rather than supplying, liquidity. The former effect dominates the latter, meaning dealers' realised spreads increase.

The sum of all these effects means that the impact on liquidity of regulation in this

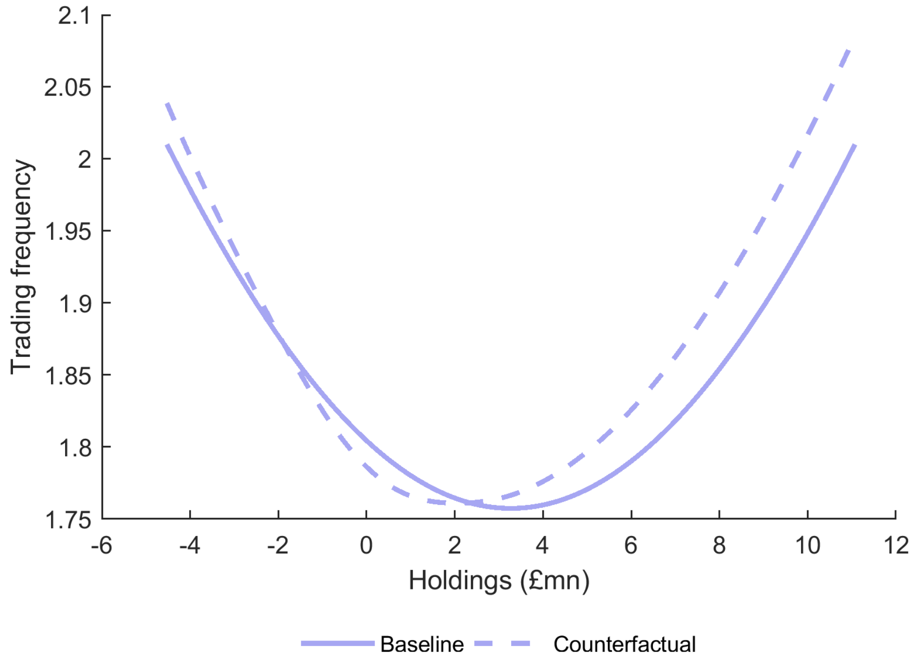


Figure 10: Dealer capital regulation and search
Note: Capital regulation is modelled as setting $\tau = 1\%$ for low-cost traders, where their flow utility from holding the asset is given by $u(\beta, h) = \beta h - 0.5\kappa h^2 - \tau|h|$.

market is negligible, and the welfare costs are limited (Table 7). Trading frequency and trade sizes increase slightly whilst waiting times and quantity rationing fall: equilibrium market functioning is no worse than without regulation. Whilst dealers' expected utility falls as their returns from holding the asset decrease, other traders benefit from regulation. This offsetting is driven by two factors. Firstly, traders compete with one another for the right to hold a bond and profit from its cashflows. A reduction in dealers' willingness to hold a bond means other traders can increase their holdings, which benefits them. Secondly, any reduction in liquidity is mitigated by the effects of endogenous search - with dealers optimally adjusting their search behaviour to minimise the harm caused by tighter regulation - and the endogenous reallocation of asset holdings from dealers to non-dealers.

7.1.1 Liquidity, Regulation and Trading in a Sell-off

The negative effects of illiquidity can be particularly large during stress, amplifying the effects of financial shocks (Brunnermeier and Pedersen, 2009). There is some evidence that liquidity during stress events – for example when a bond is downgraded or is removed from an index - deteriorated after the 2008 financial crisis, as dealers were subject to more stringent regulation (Bao, O'Hara and Zhou, 2018; Dick-Nielsen and Rossi, 2019). In each of these cases, the stress event is a shock that causes traders who are not dealers - such as traders who track market indices and insurance companies – to sell the bond. Here we set out the effects of these kinds of sell-offs by non-dealers, and how this is affected by capital regulation.

To do this, we reduce the average valuation β of high-cost traders (non-dealers) and compare the outcome to our baseline results. Specifically, we shock the high-cost traders' flow utility as follows:

$$u(\beta, h) = \beta h - \frac{1}{2}\kappa h^2 - \mu_\beta \epsilon$$

where $\epsilon \geq 0$ is the size of the shock and μ_β is the average valuation β. Table 8 summarises the effects of this. The shock to the high-cost traders leads them to sell assets, with these sales absorbed by the low-cost traders – or dealers – of whom there are fewer. The price falls. High-cost traders are worse off as a result of the shock, whilst dealers benefit from increasing their holdings of the asset at a low price.

Table 8: Impact of stress

	Percentage increase
Low-cost traders	
Expected holdings	-2.0
Expected utility	-7.9
High-cost traders	
Expected holdings	11.7
Expected utility	4.2
Aggregate	
Price	-34.5

Note: This table shows the effect of a sell-off by high-search cost traders. We decrease the mean value shock β for these traders by 5% relative to our baseline estimates, and summarise the average of key quantities above.

Figure 11 shows how the effect of these sell-offs depends on dealer regulation. In particular, we show aggregate welfare when non-dealers are subject to different levels of stress, and when dealers are subject to tighter capital regulation (blue line) and when they are not (red line). Capital regulation and stress for non-dealers both decrease welfare in bond markets,

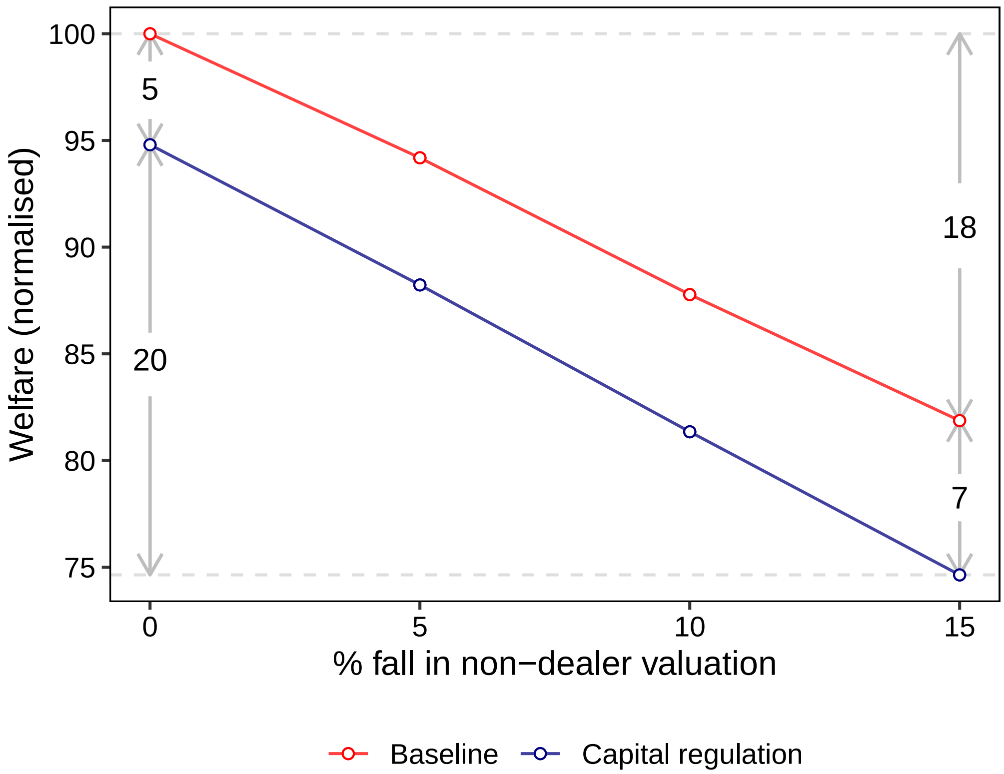


Figure 11: Effect of capital regulation in a stress

Note: This chart summarises the interaction between stress and capital regulation. Stress - along the horizontal axis - is the percentage fall in the valuation of high-cost traders (non-dealers). Specifically, we decrease high-cost traders' flow utility by $\epsilon\mu_\beta$, for $\epsilon = 0$, 5%, 10% and 15%. Dealer capital regulation is modeled as decreasing low-cost traders' flow utility by $\tau|h|$. The figures included in the chart show the welfare cost of capital with (7) and without (5) a stress, and the welfare cost of a stress with (20) and without (18) capital regulation.

and there is a positive interaction between the two. The welfare cost of the largest stress is 11% greater when dealers are subject to capital regulation. Alternatively, the welfare cost of capital regulation is 40% greater in a stress than it is in normal times.[28] This is because sell-offs of the type studied in Bao, O'Hara and Zhou (2018) and Dick-Nielsen and Rossi (2019) require dealers to be willing to buy the assets sold, but tighter regulation makes them less willing to do so, leading welfare to fall. This effect is greater when the initial shock is greater.

[28]With reference to Figure 11, the excess welfare cost of capital regulation in a stress is calculated as $100 \times (7-5)/5 = 40\%$. The excess welfare cost of stress under capital regulation is calculated as $100 \times (20 - 18)/18 = 11\%$

7.2 Trading Technologies

Trading in corporate bond markets is predominantly undertaken via traditional methods: dealers intermediate the market, and trades are organised bilaterally between traders, who often communicate on the phone. In recent years there has been a gradual increase in the use of electronic trading platforms (Anderson, Webber, Noss, Beale and Crowley-Reidy, 2015), which replace bilateral negotiation with a multilateral system where a trade quote is posted to all platform members, any of whom can bid to be counterparty to the trade. These types of platforms offer potential efficiency savings, particularly for those who don't have large client networks with which to trade. Such trading innovations require the participation of sufficient traders in order to succeed – there needs to be sufficient supply of assets and trading activity in order to make setting up such a platform worthwhile. The successful implementation of these innovations thus depends on the participation of the most efficient traders.

The take-up of these new technologies has been relatively slow (The Economist, 2020). In this section we offer an explanation for this: the most efficient traders in corporate bond markets lose out under these new technologies. To show this, we run two counterfactual simulations. First, we *reduce* and *homogenise* search costs in traders. Specifically, we take the highest value of search efficiency z in our (discretised) distribution of search costs, and counterfactually assign all traders this value of z. This captures the fact that platform-based trading has the potential to increase search efficiency, particularly for the most inefficient traders. Secondly, we study the Walrasian equilibrium, where trading is frictionless and prices are no longer bilaterally negotiated. This captures the increase in search efficiency in platforms, and also the change in trading mechanism they introduce.

Table 9 summarises the effect of these changes in search technologies on trading and welfare. The penultimate column shows the effects of reducing and homogenising traders' search costs. The traders that see their search costs decrease significantly increase their trading frequency. Those low-cost traders whose costs don't change decrease their trading frequency. The reason for this can be seen in the effect on spreads. Initially, low cost traders on average earn trading revenues by charging a spread – they sell the asset at a high price and buy it at a low price. They are able to do this because of their cost advantage over higher-cost traders. When this advantage disappears, their gains from trade decrease, and they decrease their search intensity. The loss of their cost advantage means that whilst aggregate welfare increases in the counterfactual, the welfare of the lowest-cost traders decreases. As shown in the final column of Table 9, the lowest-cost traders lose out even in the case where trading is frictionless. The introduction of these types of innovations may therefore not be in the interests of the most efficient traders.

Table 9: Counterfactual search technologies

	Baseline		Homogenous	Walrasian
	Low cost	High cost	Aggregate	Aggregate
Trades per month	1.76	0.20	1.75	0.03
Spread, bps	223	-223	0	0
Utility	16.3	13.3	14.5	14.5
Aggregate				
Price variance, bps	378		90	0
Trade size, £mn	0.5		0.5	1.9
Gross volume, £mn	0.07		0.23	0.08
Utility	13.8		14.5	14.5

Note: This table shows simulates the effect platform-based trading on trading, spreads and welfare The first two columns show results in our baseline equilibrium. The next column shows results when all traders' search costs are set to those of the lowest-cost trader. The final column shows results in the Walrasian equilibrium. Gross volume is instantaneous trading volume per trader. The spread is the difference between a trader's average selling price and buying price.

8 Conclusion

In this paper we structurally estimate a model of endogenous search in a decentralised market. Two observations lie at the heart of this model: (1) search intensity, and hence trading frequency, is an equilibrium object that responds to market conditions; and (2) liquidity supply is endogenous, trade specific, and liquidity can be and is supplied by all traders. These observations are of fundamental importance for understanding trading patterns in decentralised markets, and for evaluating the impact of policy interventions on market outcomes and welfare.

References

Admati, A. R., DeMarzo, P. M., Hellwig, M. F. and Pfleiderer, P. C. (2013), 'Fallacies, irrelevant facts, and myths in the discussion of capital regulation: Why bank equity is not socially expensive', *Max Planck Institute for Research on Collective Goods* 23.

Adrian, T., Fleming, M., Shachar, O. and Vogt, E. (2017), 'Market liquidity after the financial crisis', *Annual Review of Financial Economics* 9, 43–83.

Afonso, G. and Lagos, R. (2015), 'Trade dynamics in the market for federal funds', *Econometrica* 83(1), 263–313.

Andersen, L., Duffie, D. and Song, Y. (2019), 'Funding value adjustments', *The Journal of Finance* 74(1), 145–192.

Anderson, N., Webber, L., Noss, J., Beale, D. and Crowley-Reidy, L. (2015), 'The resilience of financial market liquidity', Bank of England Financial Stability Paper.

Asquith, P., Covert, T. and Pathak, P. (2019), 'The effects of mandatory transparency in financial market design: Evidence from the corporate bond market', NBER Working Paper 19417.

Bajaj, R., Binmore, A., Dasgupta, R. and Vo, Q.-A. (2018), 'Banks' internal capital markets: how do banks allocate capital internally?', *Bank of England Quarterly Bulletin* p. Q2.

Bank of England (2016), 'Financial stability report: November 2016', Technical report, Bank of England.

Bank of England (2021), 'Financial Policy Committee's core indicators', https://www.bankofengland.co.uk/financial-stability. Accessed: 05/08/2021.

Bao, J., O'Hara, M. and Zhou, X. A. (2018), 'The volcker rule and corporate bond market making in times of stress', *Journal of Financial Economics* 130(1), 95–113.

BCBS (2006), 'International convergence of capital measurement and capital standards', Technical report, Bank for International Settlements.

BCBS (2017), 'Basel III: finalising post-crisis reforms', Technical report, Bank for International Settlements.

Benetton, M. (2021), 'Leverage regulation and market structure: An empirical model of the UK mortgage market', *Journal of Finance* forthcoming.

Bessembinder, H., Jacobsen, S., Maxwell, W. and Venkataraman, K. (2018), 'Capital commitment and illiquidity in corporate bonds', *The Journal of Finance* 73(4), 1615–1661.

Biais, B., Declerck, F. and Moinas, S. (2016), 'Who supplies liquidity, how and when?', Bank for International Settlements Working Paper 563.

BlackRock (2015), 'Addressing market liquidity', Technical report, BlackRock.

Brancaccio, G., Li, D. and Schürhoff, N. (2020), 'Learning by trading: The case of the US market for municipal bonds', Working Paper, New York University Stern School of Business.

Brunnermeier, M. K. and Pedersen, L. H. (2009), 'Market liquidity and funding liquidity', *The review of financial studies* 22(6), 2201–2238.

Cai, Y. and Judd, K. L. (2013), 'Shape-preserving dynamic programming', *Mathematical Methods of Operations Research* 77(3), 407–421.

Chang, B. and Zhang, S. (2020), 'Endogenous market making and network formation', Working paper, London School of Economics.

Choi, J. and Huh, Y. (2018), 'Customer liquidity provision: Implications for corporate bond transaction costs', Working paper, University of Illinois Urbana-Champaign.

Coen, P. (2021), 'Information loss over the business cycle', Technical report, Toulouse School of Economics.

Coen, P. and Coen, J. (2019), 'A structural model of interbank network formation and contagion', Working paper, London School of Economics.

Czech, R., Huang, S., Lou, D. and Wang, T. (2021), 'Informed trading in government bond markets', *Journal of Financial Economics* .

Czech, R. and Roberts-Sklar, M. (2019), 'Investor behaviour and reaching for yield: Evidence from the sterling corporate bond market', *Financial Markets, Institutions & Instruments* 28(5), 347–379.

Di Maggio, M., Kermani, A. and Song, Z. (2017), 'The value of trading relations in turbulent times', *Journal of Financial Economics* 124(2), 266–284.

Diamond, D. W. and Rajan, R. G. (2001), 'Liquidity risk, liquidity creation, and financial fragility: A theory of banking', *Journal of political Economy* 109(2), 287–327.

Dick-Nielsen, J. and Rossi, M. (2019), 'The cost of immediacy for corporate bonds', *The Review of Financial Studies* 32(1), 1–41.

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

Duffie, D. (2010*a*), 'The failure mechanics of dealer banks', *Journal of Economic Perspectives* 24(1), 51–72.

Duffie, D. (2010*b*), 'Presidential address: Asset price dynamics with slow-moving capital', *The Journal of finance* 65(4), 1237–1267.

Duffie, D. (2018), 'Financial regulatory reform after the crisis: An assessment', *Management Science* 64(10), 4835–4857.

Duffie, D., Gârleanu, N. and Pedersen, L. H. (2005), 'Over-the-counter markets', *Econometrica* 73(6), 1815–1847.

Duffie, D., Gârleanu, N. and Pedersen, L. H. (2007), 'Valuation in over-the-counter markets', *The Review of Financial Studies* 20(6), 1865–1900.

Farboodi, M. (2021), 'Intermediation and voluntary exposure to counterparty risk', Working Paper, Massachusetts Institute of Technology.

Farboodi, M., Jarosch, G. and Shimer, R. (2021), 'The emergence of market structure', Working Paper, Massachusetts Institute of Technology.

Franzoni, F. A., Plazzi, A. and Çötelioğlu, E. (2019), 'What constrains liquidity provision? Evidence from hedge fund trades', CEPR discussion paper 13645.

Gârleanu, N. (2009), 'Portfolio choice and pricing in illiquid markets', *Journal of Economic Theory* 144(2), 532–564.

Gavazza, A. (2016), 'An empirical equilibrium model of a decentralized asset market', *Econometrica* 84(5), 1755–1798.

Goodman, T. (2001), 'Shape preserving interpolation by curves', Working paper, Dundee University.

Gromb, D. and Vayanos, D. (2018), 'The dynamics of financially constrained arbitrage', *The Journal of Finance* 73(4), 1713–1750.

Grossman, S. J. and Miller, M. H. (1988), 'Liquidity and market structure', *The Journal of Finance* 43(3), 617–633.

Hugonnier, J., Lester, B. and Weill, P.-O. (2020), 'Frictional intermediation in over-the-counter markets', *The Review of Economic Studies* 87(3), 1432–1469.

Kargar, M., Lester, B., Lindsay, D., Liu, S., Weill, P.-O. and Zúñiga, D. (2021), 'Corporate bond liquidity during the COVID-19 crisis', NBER Working Paper 27355.

Kashyap, A. K., Stein, J. C. and Hanson, S. (2010), 'An analysis of the impact of 'substantially heightened' capital requirements on large financial institutions', Working paper, University of Chicago Booth School of Business.

Lagos, R. and Rocheteau, G. (2007), 'Search in asset markets: Market structure, liquidity, and welfare', *American Economic Review* 97(2), 198–202.

Lagos, R. and Rocheteau, G. (2009), 'Liquidity in asset markets with search frictions', *Econometrica* 77(2), 403–426.

Liu, S. (2020), 'Dealers' search intensity in US corporate bond markets', Working paper, Tsinghua University.

Mallaburn, D., Roberts-Sklar, M. and Silvestri, L. (2019), 'Resilience of trading networks: evidence from the sterling corporate bond market', Bank of England Working Paper 813.

Modigliani, F. and Miller, M. H. (1958), 'The cost of capital, corporation finance and the theory of investment', *The American Economic Review* 1, 3.

Myers, S. C. (1977), 'Determinants of corporate borrowing', *Journal of financial economics* 5(2), 147–175.

Myers, S. C. and Majluf, N. S. (1984), 'Corporate financing and investment decisions when firms have information that investors do not have', *Journal of financial economics* 13(2), 187–221.

Neklyudov, A. (2019), 'Bid-ask spreads and the over-the-counter interdealer markets: Core and peripheral dealers', *Review of Economic Dynamics* 33, 57–84.

Robles-Garcia, C. (2020), 'Competition and incentives in mortgage markets: The role of brokers', Working paper, Stanford Graduate School of Business.

Sambalaibat, B. (2018), 'A theory of liquidity spillover between bond and CDS markets', Working paper, Princeton University.

Schultz, P. (2017), 'Inventory management by corporate bond dealers', Working paper, University of Notre Dame.

Shimer, R. and Smith, L. (2001), 'Matching, search, and heterogeneity', *The BE Journal of Macroeconomics* 1(1).

Stein, J. C. (2012), 'Monetary policy as financial stability regulation', *The Quarterly Journal of Economics* 127(1), 57–95.

The Economist (2020), 'Electronic trading platforms are challenging bond broker-dealers', Technical report.

Trebbi, F. and Xiao, K. (2017), 'Regulation and market liquidity', *Management Science* 65(5).

Üslü, S. (2019), 'Pricing and liquidity in decentralized asset markets', *Econometrica* 87(6), 2079–2140.

Vayanos, D. and Wang, J. (2013), 'Market liquidity - theory and empirical evidence', *in* 'Handbook of the Economics of Finance', Vol. 2, Elsevier, pp. 1289–1361.

Vayanos, D. and Weill, P.-O. (2008), 'A search-based theory of the on-the-run phenomenon', *The Journal of Finance* 63(3), 1361–1398.

Appendices

A1 Further Details on Data and Empirics

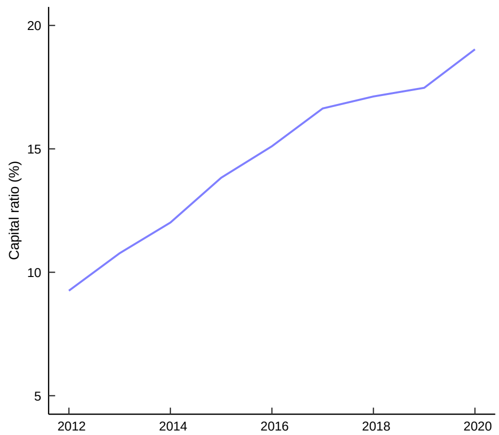
Traders are consolidated to the group level, with the exception of the asset management arms of large banking groups, who we separate from the banking business. We do this as these two business lines report separately in the data, and they tend to operate with separate balance sheets and trading strategies. We winsorize both trading prices and quantities at the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentiles. We exclude instruments that have been traded fewer than 10 times in total in our dataset, to focus on instruments where we have a meaningful number of observations.

The data present three practical challenges, common when dealing with transactions data: (a) duplicate reporting where, for example, both the buyer and the seller report the transaction; (b) how to treat agency trades, where a firm trades a bond on behalf of another firm; and (c) how to deal with trades with missing counterparty names and IDs.

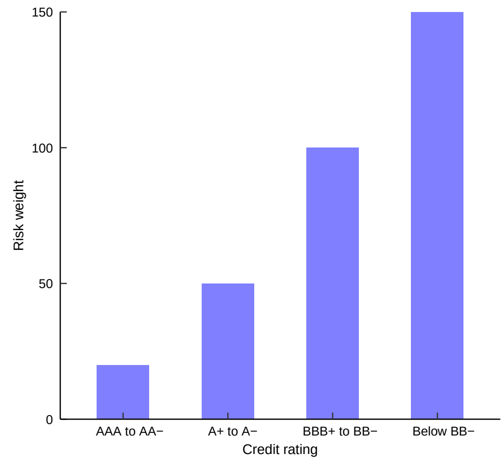
To remove duplicate trade reports, we identify all trades where the instrument, trading quantity and price match across two trade reports, and the firms reporting the trades differ. We then remove one of these duplicate reports.

With agency trades, we aim to distinguish between two types. The first is agency trading where the client firm on whose behalf the bond is being traded is a non-trading firm, for example a client of a wealth manager or a mutual fund within an asset manager's group. For the purposes of the model, we consolidate the trading firm and its clients into a single group represented by its trading entity - the wealth management firm or the asset management firm in the examples above. The second is agency trading between two trading firms. For the purposes of the model, these are trades between two distinct trading firms, and are thus not consolidated in any way. In the empirical results in Section 3 we (a) don't include agency trades when computing spreads, as agency trades don't earn a spread; and (b) treat each leg of an agency trade as a distinct trade. For example, if A buys a bond from B for C, we count this as two trades: one between A and B, and one between A and C.

In some cases a counterparty in a trade is identified only by an internal code, and not by name. This will most commonly be the case when the counterparty is not a trading firm but, for example, a client of an wealth manager. When computing firm-level summary statistics such as the number of trading firms or the distribution of trading frequencies,



(a) UK banks' capital ratios



(b) Risk weights by credit rating

Figure A1: Bank capital ratios and risk weights.

Note: These figures summarise the risk-weighted capital regulations faced by banks. The left panel shows risk-weighted capital ratios for the UK banking system, taken from Bank of England (2021). Risk weights are for exposures to corporates according to the standardised approach to risk weighting, as set out in BCBS (2006).

we do not include unnamed firms as trading firms. However, we do take into account trades with unnamed firms when computing these statistics for named firms. For example, when computing a firm's holding period for a bond, we include their trades with unnamed counterparties to accurately reflect the changes in their asset portfolios.

A1.1 Further details on institutional setting

Figure A1 shows intertemporal and cross-sectional variation in capital regulation. Figure A1a shows how UK banks' average risk-weighted capital ratios have increased through time. Figure A1b shows how capital requirements vary according to a bond's credit rating.

A1.2 Further details on Figure 2

The figure plots the empirical relationship between the likelihood of observing two offsetting trades on the same day and the theoretical relationship from a simple model with random arrivals of trades.

Let b_t denote the number of purchases of a bond made by the trader on a day, and s_t

the number of sales. We are interested in the probability of observing both a buy and a sell order for a trader on a given day, conditional on them trading on that day:

$$P(b_t \geq 1, s_t \geq 1 | b_t + s_t \geq 1) = \frac{P(b_t \geq 1, s_t \geq 1, b_t + s_t \geq 1)}{P(b_t + s_t \geq 1)} \tag{12}$$

Let trades arrive at Poisson rate λ per day, and let the trade be a buy and a sell with equal probability. Using the probability mass function of the Poisson distribution we can evaluate the probability as:

$$P(b_t \geq 1, s_t \geq 1 | b_t + s_t \geq 1) = \frac{e^{-\lambda} \sum_j^{\infty} \frac{\lambda^j}{j!} (1 - 2\left(\frac{1}{2}\right)^j}{1 - e^{-\lambda}} \tag{13}$$

A2 Computational Details

In this section we set out in detail how we solve the model.

At each step of the estimation procedure we solve for the unknown value, pricing, trading quantity, density and search intensity functions. The value shocks are assumed to take discrete values $\beta_1, ..., \beta_{n_\beta}$ and the search costs discrete values $z_1, ..., z_{n_z}$. Each of the functions are continuous in holdings h, and take the form of interpolation splines between the values $h_1, ..., h_{n_h}$. We thus need to solve for each of the functions at $n_T = n_z n_\beta n_h$ points.

We use a nonlinear solver to search over the search intensities γ and densities ϕ at each point on our grid. Conditional on these quantities, and the spline functions we fit through the holdings dimension, we can solve for the value functions V, the price function p and the quantity function q directly without resorting to numerical methods. This significantly reduces the dimensionality of the problem. Below we explain how we compute the value, pricing and quantity functions, before setting out equations we solve numerically to solve for the search and density functions.

Given our matching function (equation 8), we can solve for the value function solely in terms of the search intensities. To see this, first plug the matching function into the expression for optimal search (equation 5) and multiply by $\gamma(z, \beta, h)$, yielding:

$$\gamma(z,\beta,h)s_2(z,\gamma(z,\beta,h)) = \frac{1}{2}\iiint m(\gamma(z,\beta,h),\gamma(z',\beta',h'))S((z,\beta,h),(z',\beta',h'))\Phi(dz',d\beta',dh') \quad (14)$$

Using this equation we can substitute out the final term of the value function (equation 4):

$$rV(z,\beta,h) = u(\beta,h) - s(z,\gamma(z,\beta,h)) + \eta\int(V(z,\beta',h) - V(z,\beta,h))g(\beta')d\beta' + \gamma(z,\beta,h)s_2(z,\gamma(z,\beta,h))$$

All terms involving the value function V enter this equation linearly, the distribution function ϕ does not enter, and the functional form $s(z,\gamma)$ is known. As a result, conditional on knowing the function γ we have a closed form expression for the value function V at each of the points on our grid. Having computed the value function, we then interpolate along the holdings dimension using a cubic spline to give value functions $V(z,\beta,h)$ that are continuous in h. We can then solve exactly for trading quantity and price at each point on our grid using the Nash bargaining solution, which depends only on the value functions of the two traders that meet.

The variables we pass to the nonlinear solver thus consist of a set of search intensities γ and densities ϕ at each point on the grid. We then fit a cubic interpolation spline through the holdings grid to get a function for $\gamma(z,\beta,h)$ that is continuous in h. For the density $\phi(z,\beta,h)$ we fit a cubic Hermite spline through the holdings values, constraining the function to be a valid density - that is, taking weakly positive values and integrating to 1.[29].

The remaining equations that need solving numerically are the market clearing equation (equation **??**), the search intensity equation (equation 5) and the distribution equation (equation 6). The market clearing equation is straightforward to compute given the density function. We compute the terms of the search intensity equation based on a version of the discretized version of the density function ϕ at the points of the grid, with the surplus following directly from the value function and trading quantity.

The distribution function (equation 6) involves a double integral over the holdings of a trader and their potential counterparties involving the quantity traded, their measures and

[29] Hermite splines are particularly convenient for modelling densities as they are shape-preserving, meaning the interpolated curve and the points through which we are interpolating have the same local minima. This makes it easy to constrain the functions to be positive whilst still ensuring they are smooth. See Cai and Judd (2013) and Goodman (2001) for further details

their search intensities, which is potentially difficult to evaluate. However, a property of the Nash bargaining solution simplifies this significantly. In particular, the post-trade holdings of trader (z, β, h) after meeting another trader (z', β', h') depend only on the sum of the two traders' pre-trade holdings, $h^T \equiv h + h'$. We can thus define $h^1(z, \beta, z', \beta', h^T)$ as the post-trade holdings of trader (z, β, h) after meeting trader $(z', \beta', h^T - h)$. Fitting a cubic spline through the holdings dimension of $h^1()$, for any given h^* we can solve for the level of h^T such that $h^1(z, \beta, z', \beta', h^T) = h^*$.

The flows into and out of holdings $h < h^*$ from trades between types (z, β) and (z', β') are then as shown in Figure A2, and to get the trading flows we simply integrate meeting rates over the relevant areas. Given we have splines for the relevant expressions, this is an analytical integral, and is straightforward to compute. This enables us to compute the terms of the distribution function (equation 6).

This process enables us to solve the model at each step of the estimation procedure. We then compute the moments set out in Appendix A3 based on the model solution, and choose the parameters to minimise the distance between these moments and their theoretical counterparts.

A3 Moments

In this section we provided further details on some of the moments used in estimation.

- Correlation between quantity sold and holdings, within-traders.

$$corr^W(h, q) = \int \frac{cov(h, q|z)}{\sqrt{\mathbb{V}(h|z)\mathbb{V}(q|z)}} f(z) dz$$

where:

$$cov(h, q|z) = \iint \frac{\gamma(z, \beta, h)\phi(z, \beta, h)}{\Gamma(z) f(z)} \times \int \big(q(z, \beta, h, \Delta') - \mathbb{E}(q|z)\big)\big(h - s\big) \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\beta dh$$

$$\mathbb{V}_{TW}(h|z) = \iint \frac{\gamma(z, \beta, h)\phi(z, \beta, h)}{\Gamma(z) f(z)} \int \big(h - s\big)^2 \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\beta dh$$

$$\mathbb{V}(q|z) = \iint \frac{\gamma(z, \beta, h)\phi(z, \beta, h)}{\Gamma(z) f(z)} \int \big(q(z, \beta, h, \Delta') - \mathbb{E}(q|z)\big)^2 \frac{\gamma(\Delta')\phi(\Delta')}{\Gamma} d\Delta' d\beta dh$$

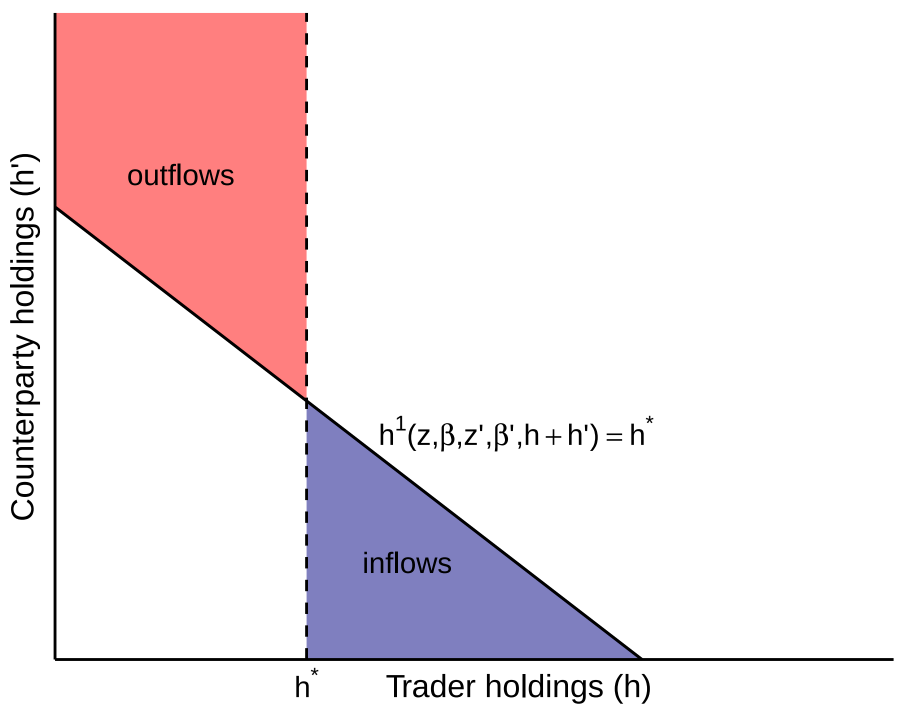


Figure A2: Trading inflows to and outflows from $h \leq h^*$

Note: This figure shows the inflows to and outflows from holdings $h < h^*$. For given (z, β, z', β'), if a trader and counterparty meet with pre-trade holdings in the blue shaded area, the trader's post-trade holdings will be beneath h^*, representing an inflow. If their holdings are in the red shaded area, the trader's post-trade holdings will be above h^*, representing an outflow. The solid line, with gradient -1, denotes meetings that will result in the trader having post-trade holdings of exactly h^*.

- Correlation between absolute inventory $inv \equiv |h - s|$ and trading frequency, within traders.

$$corr^W(inv, s) = \int \frac{cov(inv, n|z)}{\sqrt{\mathbb{V}(inv|z)\mathbb{V}(n|z)}} f(z) dz$$

where:

$$\mathbb{V}(inv|z) = \iint inv^2 \frac{\phi(z, \beta, h)}{f(z)} d\beta dh$$

$$cov(inv, n|z) = \iint \big(2\gamma(z, \beta, h)\pi(z, \beta, h) - n(z)\big) inv \frac{\phi(z, \beta, h)}{f(z)} d\beta dh$$

$$\mathbb{V}(n|z) = \iint \big(2\gamma(z, \beta, h)\pi(z, \beta, h) - n(z)\big)^2 \frac{\phi(z, \beta, h)}{f(z)} d\beta dh$$