Orderbook spreads and depth in a dynamic general equilibrium model

Scott Condie* Brennan Platt†

ABSTRACT

This paper studies equilibrium orderbook formation in a limit-order market. We model the market using a specialized search-theoretic model where the shape of the orderbook and its spread are determined jointly in equilibrium. The model characterizes liquidity as a function of differences in valuation between sellers and buyers, beliefs about the probability distribution of the arrival of buyers and sellers, exchange fees and preference parameters like patience and beliefs about the expected lifetime of information. The efficiency of the market is characterized as a function of these parameters. Additionally, we demonstrate how to extract the model's parameters from orderbook data and, using a sample of data from Coinbase's Bitcoin/U.S. Dollar exchange, characterize the extent to which changes in market liquidity are determined by each of these factors. This model and the empirical tools developed to estimate it are useful to study the effects of changes in market fees, trader arrival rates, and heterogeneous information on market liquidity as measured by both the spread and price impact.

Keywords: Orderbook, spread, price impact, equilibrium search

JEL classification: G14, D53

1 Introduction

In most equity markets throughout the world, market makers provide liquidity to those wishing to purchase or sell assets by placing limit orders that remain on the book until either executed or canceled. The liquidity of these markets is often characterized in two specific ways: spread and price impact. The spread is the difference between the highest bid

*Department of Economics, Brigham Young University. Email: ssc@byu.edu. Telephone: +1.801.422.5306.

†Department of Economics, Brigham Young University. Email: bcplatt@byu.edu. Telephone +1.801.422.8904.

(or buy) limit order and the lowest ask (or sell) limit order. The spread represents the cost associated with obtaining immediate execution of a trader's desired position. More liquid markets are characterized by narrower spreads.

The price impact in a limit order market is the extent to which the price moves as greater quantities of shares are purchased or sold. One market is more liquid than another if its price impact is smaller. That is, if purchasing 100 shares in market A leads to a change in price of \$0.50 between the price at which the first share is purchased and the price at which the 100th share is purchased and in market B that price change is only \$0.20, then market B is more liquid than market A.

As a market's liquidity increases, the market has greater informational and transactional efficiency. There is evidence that a lack of liquidity is perceived by traders to be risky in the aggregate (Pástor & Stambaugh 2003), causes higher fees associated with seasoned equity offerings (Butler et al. 2005), and leads to less informative prices (Kerr et al. 2020). These issues suggest that understanding the factors that determine liquidity in equilibrium is an important area of study.

This paper presents a model of liquidity that allows for an equilibrium characterization of both spreads and price impact. It does this without the presence of noise traders; that is, rather than having some traders who must transact regardless of price, the preferences of all participants are fully modeled. This allows us to study how changes in market characteristics (*e.g.* the relative quantity of market makers, the fees associated with providing or removing liquidity, or differences in opinion about the value of the asset) affect the equilibrium prices in the orderbook and hence measures of market liquidity. This equilibrium model also enables meaningful evaluation of welfare gains associated with the market.

The model is based on a dynamic search model where buyers arrive at the market looking to immediately purchase an asset for which limit offers have been established.[1] Buyers'

[1]For ease of exposition, we focus only on the ask side of the market, though the bid side would proceed

valuations of the asset imply that for at least some offered prices they are willing to purchase the asset, although they are not willing to purchase the asset for any offered price. If they do not purchase the asset, they wait for the next period to see if a more attractive offer is available to them.

Sellers in the market are liquidity providers and place limit orders to sell the asset based on their subjective valuation of the asset. They place limit orders at prices that are optimal given the publicly-known demand parameters as well as the understanding that higher-priced limit orders will receive lower priority than lower-priced limit orders.

In this model we characterize steady-state orderbooks, or distributions of offers. Indeed, prices are dispersed continuously, even though all buyers share a common view of the asset's fundamental value. We omit differing valuations from the model to highlight the role of the orderbook itself in creating its own dispersion. Variation in the flow of market orders mean that the probability a limit order being executed decreases in the price of that order (for the ask side of the orderbook), setting up a trade off for sellers between higher markups versus a longer wait to transact. We explore how much of orderbook prices can be explained from variation in participant arrivals rather than variation in participant valuations.

This characterization allows us to answer equilibrium questions which are more difficult to answer in partial equilibrium frameworks. We show, for example, that increases in the fee charged to liquidity takers leads to liquidity providers placing more attractive offers and hence more liquid markets *ceteris paribus*. Intuitively, this occurs because, holding the internal valuation of the stock constant for both buyers and sellers, an increase in the take fee makes the asset less desirable for buyers which increases their bargaining power relative to sellers. We also find that increases in the popularity or salience of a particular stock to buyers has mixed results on the liquidity in the orderbook. The increase in popularity decreases spreads, but increases equilibrium price impact. Furthermore, we show that increases in the

similarly.

difference between buyers' and sellers' valuations of the asset lead to less liquid markets. This happens because as buyers' desire for the asset increases relative to sellers', the sellers' bargaining position increases and they can sell at more favorable prices for them.

To illustrate its potential explanatory power, we apply the model to orderbook data from Coinbase's Bitcoin/U.S. Dollar exchange. Our calibration process interprets the data through the lens of our model. First, we parse the flow of updates to the orderbook into time periods (regimes) in which the best ask and best bid prices do not cross during the regime. Then we provide an estimation technique that estimates the average orderbook in each regime, while accounting for bias that naturally arises when the lowest price is executed first and thus is frequently unobserved. We can then extract the spread and price impact from the data and use these to estimate the market friction and distribution of participant arrivals. Compared across regimes, this process sheds light on the underlying causes of shifting orderbook prices.

With buyer and seller preferences modeled (rather than relying on noise traders), the costs and benefits to all market participants can be quantified in welfare calculations. The theory implies that potential welfare gains are more fully realized as the flows of buyers and sellers are closer to equal. We apply our estimation of the market frictions to evaluate the efficiency of trading on the Coinbase BTC/USD exchange, and examine whether spread or price impact are better signals of inefficiency.

We proceed in the paper by first reviewing the literature, after which we build a general model and discuss its equilibrium in section 2. In section 3 we illustrate how to apply the general model to observable orderbook data. Section 4 discusses the welfare implications of orderbooks, while section 5 concludes.

1.1 Literature

Two standard models of the orderbook are Glosten & Milgrom (1985) and Kyle (1985). Glosten & Milgrom (1985) models market makers as submitting limit orders so as to reflect the information content in the arriving order being a buy or sell. The arrival process of market orders is given exogenously and the best bid and offer prices are then derived optimally in a competitive environment. Our paper differs from Glosten & Milgrom (1985) in that we model the preferences of both liquidity demanders and liquidity suppliers. In so doing, we can trace out the shape of the order book, as opposed to describing just the best bid and offer. In modeling both liquidity suppliers and demanders we can say more about each side of the market's motivations and estimate welfare effeccts, at the cost of increased complexity.

Kyle (1985) models a game of demand curve submission that can be interpreted as a primitive to forming the orderbook. Like in Glosten & Milgrom (1985), liquidity demanders are exogenous and divided between informed and uninformed traders. However, informed traders in the Kyle model understand the impact that their orderflow will have on the price at which their orders are executed and so they shade their bids to maximize the expected profit from their trades. This model becomes tractable through a set of parametric assumptions (linearity and normality). In our model, trade is motivated based on disagreements about asset values (rather than differential information), and liquidity suppliers submit limit orders (rather than supply/demand curves).

Back & Baruch (2004) establishes a connection between Glosten & Milgrom (1985) and Kyle (1985) by nesting each of these in a continuous time model and establishing conditions under which the equilibrium of their model converges to that of Glosten & Milgrom (1985) and Kyle (1985). Their analysis focuses on (for Glosten-Milgrom) characterizing optimal trade timing for a single informed trader, and showing conditions under which this version of the model will converge to the equilibrium in Kyle. Although we discuss some of the connections between our results and those of Glosten & Milgrom (1985) and Kyle (1985), our modeling framework is sufficiently different from those models that we do not attempt to study conditions under which our model would converge to those.

Goettler et al. (2005) build a dynamic model of limit order markets where all traders' preferences are fully modeled, as in our paper. In that paper, traders arrive with random valuations and liquidity needs and must decide whether to buy or sell the asset. Information is disseminated over time in such a way that those who have previously submitted orders on the book did so with information that becomes outdated over time. Orders are cancelled exogenously when certain market conditions are met. Goettler et al. (2005) differs in several key ways from the model presented in this paper. First, in their model traders arrive at a constant rate and the orderbook is built up over time, while in ours traders arrive according to a Poisson rate given exogenously and we characterize the steady state orderbook. In Goettler et al. (2005) information changes over time in ways that lead to accumulation of orders based on changes in that information and the properties of the exogenous order cancellation function, while in our model orders are placed and executed endogenously. Goettler et al. (2005) allows for traders to endogenously determine whether they will buy or sell and whether they will do so with limit orders or market orders, where our model designates traders as market makers or liquidity demanders exogenously. Finally, Goettler et al. (2005) must be solved numerically, while our model has closed form solutions.

Foucault et al. (2005) build a model of the orderbook that relies on differences in the patience of market participants, with patient traders submitting limit orders and impatient traders submitting market orders. Traders arrive according to a Poisson process and buyers and sellers have exogenously given differences in valuations, as in this paper. Unlike in this paper, Foucault et al. (2005) assume that all limit orders that are submitted must be price improving (i.e. narrow the spread) and that buyers and sellers alternate with certainty. These assumptions (which can be relaxed somewhat, as they show) allow them to study the expected waiting time for a limit order to be fulfilled. Unlike that paper, our model obtains a closed form characterization of both depth and spread (in steady state), whereas they characterize a distribution of spreads and extensively study executions times and market dynamics.

Several search theoretic papers study liquidity in asset markets from a macro perspective. Lagos & Rocheteau (2009) studies liquidity in over-the-counter markets. This model makes predictions about liquidity measures like the bid-ask spread as in our paper, although they do not attempt to model the mechanics of price priority in limit orderbooks as is done here. Cui & Radde (2016) also abstracts from the specific properties of the limit orderbook, embedding a financial sector with search frictions into a dynamic general equilibrium consumption-saving-investment model. Vayanos & Wang (2007) builds a search-based model of trading with the friction that traders can search for only one asset and then studies the equilibria that result. They show that in one equilibrium, short-horizon investors congregate in one market and that this market is more liquid than the other market with long-horizon investors. The question of time horizon does not enter into our model since all traders have the same trading horizon.

2 An orderbook model

In this paper we will explicitly model the ask (or offer) side of the orderbook. The bid side can be modelled analogously. The model is populated by a competitive mass of sellers, who place limit orders on the orderbook, and a flow of buyers who arrive randomly and place market orders to purchase the asset.[2] Buyers always transact with the lowest priced order

[2]It will be assumed later that buyers have an endogenously-determined maximum price they are willing to pay this period, above which they would prefer to wait for next period before submitting an order. This behavior is consistent with their submission of marketable limit orders–or, limit buy orders that have a price that crosses the spread. Throughought the paper we use the term "market orders" as a short hand for these marketable limit orders.

on the orderbook, but if multiple market orders arrive simultaneously, the order of execution is randomized, creating ex-ante uncertainty about the realized transaction price for a given buyer.

2.1 Sellers

Each seller places a limit order to sell one unit of the asset. Sellers see the relative flow value of holding the asset as d_A. A seller's ask at price a will transact with endogenous probability $G(a)$ over a unit of time, upon which the exchange will provide the seller a rebate f_m for providing liquidity. Let $V_S(a)$ denote the seller's present expected value of placing a limit order with price a. At Poisson rate θ, the trading opportunities cease and the seller continues to receive flow value d_A indefinitely.[3] This trader's Bellman function is thus:

$$\rho V_S(a) = d_A + G(a)(a + f_m - V_S(a)) - \theta \left(\frac{d_A}{\rho} - V_S(a) \right), \tag{1}$$

which implies that

$$V_S(a) = \frac{(\rho + \theta) d_A + \rho G(a)(a + f_m)}{\rho(\theta + \rho + G(a))}. \tag{2}$$

Note that the rate of time preference, ρ, and the rate of market obsolescence, θ, have similar roles in discounting future opportunities whether they are farther away or less likely to occur. Thus, we can call their sum, $\rho + \theta$, the effective discount rate.

Since sellers are risk-neutral and competitive, they must be indifferent between any prices

[3] This shock occurs to all buyers and sellers simultaneously. This shock can be interpreted as buyers' information becoming obsolete, removing their interest in buying the asset. One could further model this as a shock that initiates a new steady state, with transitions between steady state regimes depicted as a Markov process, but this would have minimal impact on the model predictions with significant notational complication.

offered in equilibrium. This implies that for any price a in the support of the orderbook,

$$\frac{\partial V_S(a)}{\partial a} = 0 \quad \Longrightarrow \quad G'(a) = \frac{\rho(\rho + \theta + G(a))}{(\rho + \theta)(d_A - \rho\,(a + f_m))} G(a) \tag{3}$$

If we assume that the lowest price a_0 will transact with probability 1 each unit of time, we can solve this differential equation for $G(a)$ to obtain

$$G(a) = \frac{\rho\,(a_0 + f_m) - d_A}{\rho\left(\frac{a - a_0}{\rho + \theta} + f_m + a\right) - d_A}. \tag{4}$$

Given this solution for $G(a)$, we can solve for the value function to get

$$V_S(a) = \frac{d_A}{\rho} + \frac{a_0 + f_m - \frac{d_A}{\rho}}{1 + \rho + \theta}. \tag{5}$$

This value function depends on the endogenous value of a_0 which will be determined in what follows.

2.2 Market Execution

One key feature of the order book is its order of execution. Each market order is first crossed with the lowest available limit order, moving upward if there is not sufficient liquidity at the initial price to cover the order. Therefore, the probability $G(a)$ of a limit order asking price of a being executed depends on the number of buyers who arrive. Define q as the ratio at any moment of shares desired to be purchased by buyers over shares offered by sellers on the orderbook. We assume this ratio is randomly distributed according to an exogenous cumulative distribution $F(q)$.

As an example of such a distribution, suppose that the number of shares n_a that sellers offer as limit orders are exponentially distributed with mean $1/\alpha$, while the number of shares

n_b that buyers seek as market orders are independently exponentially distributed with mean $1/\beta$. The joint distribution then becomes $f(n_a, n_b) = \alpha e^{-\alpha n_a} \beta e^{-\beta n_b}$. If we define the ratio of market orders to limit orders as $q = \frac{n_b}{n_a}$ and define $\phi = \frac{\beta}{\alpha}$ (the average number of sellers per buyer), this cumulative density can be transformed to become $F(q) = \frac{\phi q}{1+\phi q}$.

If the realized ratio has $q < 1$, then not all limit orders will be executed in equilibrium. If instead $q \geq 1$, then all of the limit orders on the book (below a price to be determined) will be executed, while some market orders will not be executed. Therefore, in equilibrium under rational expectations prices a and realized participation q must be related by the equation:

$$G(a) = 1 - F(q). \tag{6}$$

That is, beliefs about the likelihood of a limit order being executed should depend on the distribution of buyers relative to sellers. Substituting equation (4) into equation (6) and solving for a gives the optimal limit order price as a function of the realized ratio q of buyers to sellers who arrive in the market:

$$a = a_0 + \left(a_0 + f_m - \frac{d_A}{\rho}\right)\left(\frac{\rho+\theta}{1+\rho+\theta}\right)\left(\frac{F(q)}{1-F(q)}\right). \tag{7}$$

From this it can be seen that in equilibrium, each price a is a function of the (endogenous) minimum bid a_0, the difference between the revenue from a sale at the minimum price and the lifetime value of a seller holding the asset ($a_0 + f_m - {}^{d_A}/_{\rho}$), which in equilibrium will be weakly positive (since sellers would not place an offer that has a lower value than their lifetime value of holding the asset, the effective discount rate $\rho + \theta$, and the cumulative hazard function ${}^{F(q)}/_{(1-F(q))}$ of the ratio q. This cumulative hazard function is increasing in q, indicating that the price a is increasing in q, as expected.

In equilibrium, the highest ask price that could be executed corresponds to $q = 1$, since for all $q > 1$ all asks will be executed and some buyers will not desire to purchase the asset. Thus, we can find a_1, the maximum offered price, to be

$$a_1 = a_0 + \left(a_0 + f_m - \frac{d_A}{\rho}\right)\left(\frac{\rho+\theta}{1+\rho+\theta}\right)\left(\frac{F(1)}{1-F(1)}\right). \tag{8}$$

The probability of the price a_1 transacting is given by

$$G(a_1) = 1 - F(1). \tag{9}$$

2.3 Buyers

We now turn to the buyers' side of the market. Buyers are assumed to receive relative flow value d_B from holding the asset, where we assume that $d_B > d_A$ to ensure that the buyers who arrive want to purchase the asset from sellers.[4] If they successfully purchase the asset, buyers must also pay the additional take fee f_t. If buyers do not successfully purchase the asset below their maximimum willingness to pay, they continue to search in the next period. At rate θ, the trader's information about the asset become obsolete and the buyer's search ends.

Buyers' outcomes can fall into one of three categories. If the ratio of buyers to sellers q satisfies $q \leq 1$, then all buyers get to transact (at various prices). If $q > 1$, more buyers arrive than there are shares available to sell, so a buyer can either be one of those lucky enough to purchase the asset, or they may be in the set of buyers who can't purchase the asset that period.

[4]The difference in flow value between buyers and sellers is fundamental to generating trade. In the analogous bid market, if modeled, those placing limit orders to buy must value the asset more than those placing market orders to sell, but both would have to value the asset less than participants in the ask market.

Given these three possible outcomes, the value function for a buyer is

$$\begin{aligned}\rho W_B = &\int_0^1 \left[\int_0^q \frac{1}{q} \left(\frac{d_B}{\rho} - a(s) - f_t - W_B \right) ds \right] F'(q) dq \\ &+ \int_1^\infty \left(\int_0^1 \left(\frac{d_B}{\rho} - a(s) - f_t - W_B \right) ds \right) \frac{1}{q} F'(q) dq + \theta(0 - W_B).\end{aligned} \tag{10}$$

The first term in this expression represents the change in the value that comes from successfully transacting when $q \leq 1$. Notice that while the buyer knows she will transact if $q \leq 1$, she doesn't know what her placement in the order arrival process will be, so the price she will receive is uniformly distributed over the orderbook. The second term in this expression represents the expected flow value to the buyer when $q > 1$, where there is some probability that the buyer won't transact at all, and if she does transact then it will be a price that is uniformly distributed over the prices on the orderbook.

After substituting in the expression for the price given in equation (7), this becomes:

$$W_B = \frac{d_B}{\rho} - (a_0 + f_t) - \frac{\rho + \theta}{\rho + \theta + X_t} \left(\frac{X_m}{1 + \rho + \theta} \left(a_0 + f_m - \frac{d_A}{\rho} \right) + \frac{d_B}{\rho} - a_0 + f_t \right) \tag{11}$$

where

$$\begin{aligned}X_t &= F(1) + \int_1^\infty \frac{1}{q} F'(q) dq \\ X_m &= \int_1^\infty \frac{1}{q} F'(q) \left(\int_0^1 \frac{F(s)}{1 - F(s)} ds \right) dq + \int_0^1 \frac{1}{q} F'(q) \left(\int_0^q \frac{F(s)}{1 - F(s)} ds \right) dq.\end{aligned} \tag{12}$$

Here, X_t computes the probability that liquidity takers (i.e. buyers) successfully transact and X_m is related to the probability that liquidity makers (i.e. sellers) successfully transact. For the exponential distribution introduced in the previous section, these evaluate to $X_t = \phi \log\left(1 + \frac{1}{\phi}\right)$ and $X_m = \frac{1}{\phi} \log(1 + \phi)$.

In order to find the maximum amount the buyer would be willing to pay, and hence the price at which they would place their marketable limit order, we find the point at which the buyer is indifferent between a price a on the book and waiting to purchase the asset. This price is the solution a_1 to

$$\frac{d_B}{\rho} - (a_1 + f_t) = W_B. \tag{13}$$

Combining Equations (8), (11) and (13) allows us to solve for the best ask that will be placed in equilibrium. That is, we are requiring that the maximum willingness to pay of buyers be consistent with the maximum price that sellers expect to be able to charge.

$$a_0 = \frac{d_A}{\rho} - f_m + \frac{(1+\rho+\theta)}{\frac{\rho+\theta+X_t}{1-F(1)} + 1 - X_m - X_t}\left(\frac{d_B - d_A}{\rho} - f_t + f_m\right). \tag{14}$$

The denominator in Equation (14) appears frequently, which we define as:

$$L \equiv \frac{\rho+\theta+X_t}{1-F(1)} + 1 - X_m - X_t \tag{15}$$

The term L reflects the underlying liquidity of the market, which we explore in the next subsection. Using the solution for a_0, we can substitute into equation (7) to get the equilibrium orderbook relationship between the transaction price if fraction $q \leq 1$ of the available shares are transacted.

$$\begin{aligned} a(q) &= \frac{d_A}{\rho} - f_m + \frac{\left(\frac{d_B-d_A}{\rho} - f_t + f_m\right)}{L}\left(1+\rho+\theta+\frac{(\rho+\theta)F(q)}{1-F(q)}\right) \\ &= \frac{d_A}{\rho} - f_m + \frac{\Delta V}{L}\left(1 + \frac{\rho+\theta}{1-F(q)}\right) \end{aligned} \tag{16}$$

where $\Delta V \equiv \frac{d_B-d_A}{\rho} - f_t + f_m$ is the difference in net valuations between buyers and sellers. Note that the orderbook can be constructed from the price function $a(q)$ on the horizontal axis and q times the orderbook volume on the vertical axis. With our example of exponentially distributed arrivals, $a(q)$ is a linear function of q, but in general, $a(q)$ has the same shape as $\frac{1}{1-F(q)}$. Notice that if there were Bertrand competition among sellers, each would receive $\frac{d_A}{\rho} - f_m$ for their asset, exactly compensating them for their lowest willingness to sell. As such, the term

$$\frac{\Delta V}{L}\left(1 + \frac{\rho + \theta}{1 - F(q)}\right) \tag{17}$$

depicts how trading in an orderbook market (with its inherent frictions of unbalanced arrivals captured in $F(q)$) will raise the price above this competitive outcome. These frictions involve the difference in net valuations between buyers and sellers, ΔV; the effective discount rate $\rho + \theta$, which characterizes the cost of waiting another period to match; and the distribution of buyers and sellers, given in $F(\cdot)$ (which also appears in X_m and X_t).

2.4 Market liquidity

In limit order markets, the liquidity of an asset is often defined by the spread of the asset and the price impact. From equation (16), we can calculate the spread and price impact in this market as a function of the model parameters. We define the spread to be

$$\begin{aligned} S &= a_0 - \left(\frac{d_A}{\rho} - f_m\right) \\ &= (1 + \rho + \theta)\frac{\Delta V}{L} \end{aligned} \tag{18}$$

which is the difference between the best price that the orderbook ever offers, a_0, and the price that would come from a competitive market.

The price impact for an order that absorbs the fraction q of the orderbooks in our model is defined to be

$$
\begin{aligned}
PI(q) &= a(q) - a_0 \\
&= \left(\left(\frac{d_B}{\rho} - f_t\right) - \left(\frac{d_A}{\rho} - f_m\right)\right) \frac{\rho+\theta}{\frac{\rho+\theta+X_t}{1-F(1)} + 1 - X_m - X_t} \frac{F(q)}{1-F(q)} \\
&= (\rho+\theta)\frac{\Delta V}{L}\frac{F(q)}{1-F(q)}.
\end{aligned}
\tag{19}
$$

Given the similar terms involved in each, it is natural to consider the ratio of price impact to spread, which is:

$$
\frac{PI(q)}{S} = \frac{\rho+\theta}{1+\rho+\theta} \cdot \frac{F(q)}{1-F(q)} \tag{20}
$$

These expressions show that any increase in ΔV, such as higher buyer valuations, lower seller valuations, or an increased spread in transaction fees, will proportionally affect both measures of market illiquidity, yet leave the ratio PI/S unchanged. If the distribution of participant arrivals $F(q)$ were to increase at all q (i.e. to a first-order stochastically dominated distribution), the ratio PI/S will also rise. For instance, in our exponential arrival example for $F(q)$, an increase in ϕ causes $F(q)$ to rise; we illustrate such a change in Figure 1.

This expression shows that market liquidity is a function of differences in beliefs between buyers and sellers, the metric of demand/supply size similarity in L and the distribution of arrival of buyers and sellers. For the spread, the distribution $F(\cdot)$ enters only through the liquidity measure L. For price impact, the distribution $F(\cdot)$ enters directly. Spreads and price impact are also increasing in $\rho+\theta$. Regimes where traders are impatient or information becomes obsolete quickly will be less liquid in equilibrium.

These equilibrium characterizations of market liquidity demonstrate several important factors in understanding variations in liquidity across exchanges or assets. Differences in valuations between buyers and sellers can arise through tailored marketing campaigns, substantive differences in the interpretation of available public data, differences in information

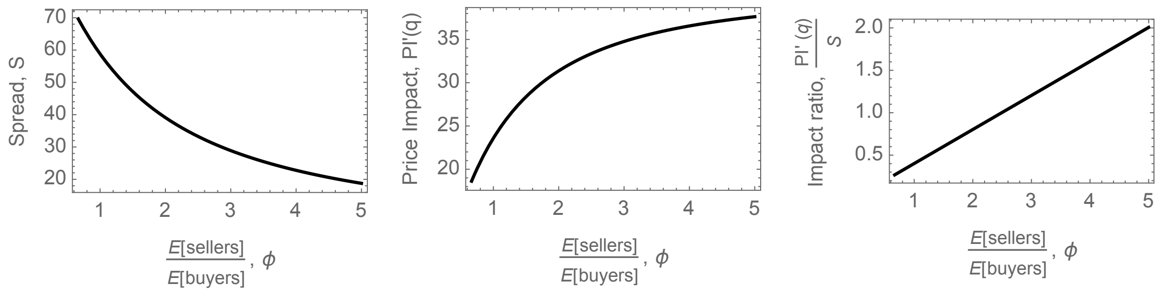


Figure 1: The equilibrium spread, price impact, and their ratio as a function of ϕ

gathering technologies or preferences amongst market participants, differences in hedging demands from alternative portfolio goals, or alternative tax or regulatory constraints among market participants. In situations where this heterogeneity is more extreme, we would expect markets to be less liquid.

Additionally, ΔV depends on the fee spread $f_m - f_t$ provided to market participants by the exchange. Exchanges where the fee spread is high will be less liquid, *ceteris paribus*. This result differs from the result presented in Colliard & Foucault (2012) in that here the effect on market liquidity from decreases in the revenue collected by the exchange unambiguously increase the equilibrium liquidity in the market.

Additionally, equilibrium liquidity depends on the nature of the arrival process of traders in two ways. The first relates to the relative balance of buyers and sellers characterized in L. It is useful to note that the liquidity measures presented here need not move in tandem. For the parametric example considered in this section, imagine a set of information factors that decrease ϕ. These might include extensive marketing, "pump and dump" schemes, or the narrow release of insider information.

Proposition 1. *The spread and price impact will decrease with:*

1. *a smaller difference between buyers and sellers valuations, ($d_B - d_A$, in ΔV)*

2. *a smaller fee spread is small ($f_m - f_t$ in ΔV)*

3. larger aggregate liquidity L

4. more patient traders or longer-lasting information (smaller ρ or θ).

The first three have proportionate impact on both measures, holding the ratio of $PI(q)/S$ constant, while reductions in effective discount rate will lower the ratio.

2.5 Comparison with the Glosten & Milgrom (1985) and Kyle (1985) models

One of the key differences between the model in this paper and the workhorse model of Glosten & Milgrom (1985) and Kyle (1985) is the absence of noise traders in the present model. This absence means that variables like the fraction of traders who are believed to be informed (as in Glosten & Milgrom (1985)) or the relative size of noise trader demand relative to fundamental uncertainty (as in Kyle (1985)) have no analog in our paper. However, the issue of spread size addressed in Glosten & Milgrom (1985) and price impact, addressed in Kyle (1985), can be compared.

Several key pieces of intuition about the formulation of the spread arise from Glosten & Milgrom (1985). These include the fact that competition amongst liquidity providers narrows spreads and that increases in the fraction of traders who are thought to be informed increases spreads. We see a similar result in terms of spread. An increase in the competitiveness of the offer side (as parameterized by an increase in ϕ in the example), leads to narrower spreads. Glosten & Milgrom (1985) also contains the result that increases in the variability of the underlying asset value (as characterized by the difference in the value of the asset in the good state of the world vs. the bad state of the world) lead to increases in the spread. While not directly analogous, in our model, differences in the valuation of market makers vs. liquidity demanders also lead to increases in the spread and a reduction in market liquidity.

Kyle (1985) characterizes the equilibrium price in the market as an increasing function of the ratio of the underlying asset variation relative to variance of noise trader demand. If one interprets the difference in values between buyers and sellers in our model as being positively related to the underlying fundamental uncertainty about asset values, then this model and Kyle's model give results that are consistent in that price impact is increasing in this measure. Our model goes further in relating more competition among buyers to a larger price impact.

3 Estimating model parameters with orderbook data

In this section, we apply the model to orderbook data from Coinbase's Bitcoin/USD exchange, which we parse into 2480 segments of time (which we call *regimes*). We first describe key features of the data, then develop a method for extracting analogs of the model parameters from the data. We then analyze each of the regimes, first by inspecting four representative examples, then by describing the aggregate fit of the model. This allows the model to decompose the underlying factors driving the observed variation in prices and liquidity across regimes.

3.1 Data Description

The Coinbase Bitcoin/USD exchange provides a live feed of orderbook updates through their Websocket API. These updates provide all changes to the orderbook in near real-time. In order to demonstrate the technique used here, we limit our analysis to one hour,[5] which yielded 1,321,084 orderbook updates and 14,807 order executions. These updates occur at the rate of approximately 353 updates per second and 3.95 executions per second.

Our steady-state model is most applicable during periods when the orderbook is relatively

[5] Data was collected on June 20, 2023, from approximately 22:25:51.04 UTC to 23:28:18.46 UTC.

Table 1: Regime summary statistics

Regimes ($N = 2480$)	Duration (sec)	Updates (#)
Mean	1.51	533
St. Dev.	3.07	836
Min	0.00	2
25%	0.04	97
50%	0.35	244
75%	1.60	664
Max	35.93	12820

stable for long enough that participants can fully adjust to market conditions. In settings with algorithmic traders, that stability can plausibly be achieved in seconds rather than (as in macro models) months or years. To identify potential steady state regimes, we begin a regime at the start of trading and track the highest best bid and the lowest best ask over time. We end a regime and start a new one when the best ask and bid cross, repeating this process until the end of the sample. In so doing, we generate 2480 regimes. Some of these regimes are very short, lasting only microseconds, while others last for several seconds.

Table 1 gives statistics on the duration and the number of updates for these regimes. Not surprisingly, the duration of the regimes varies. The median regime lasts about 0.35 seconds. The longest steady-state regime lasts about 36 seconds and the mean is 1.51 seconds. The median number of orderbook updates performed during a regime is 244 updates, with a mean of 533 updates and a maximum of 12,820 updates.[6] We anticipate that our model is most applicable to regimes with longer durations; those with extremely short durations occur when both bid and ask prices are rapidly rising or falling, which our model would interpret as a change in other fundamental parameters.

Next we report price statistics from the data period. Figure 2 shows the evolution of

[6]Note that the duration quartiles do not necessarily correspond to the same regime as the corresponding quartile in the distribution of number of updates, since the rate of updates per second is not constant.

Table 2: Summary statistics of Price, Spread and Price Impact for Averaged Orderbooks

($N = 2480$)	Price	Spread	Price Impact
Mean	\$28,218.92	\$0.16	\$2.12
St. Dev.	\$23.92	\$0.38	\$1.48
Min	\$28,167.07	\$0.01	\$0.00
Median	\$28,219.10	\$0.01	\$1.81
Max	\$28,284.36	\$4.23	\$9.71

the best ask (price), the spread, and the price impact (computed for a purchase of 2 BTC) for the steady-state averaged orderbooks, while Table 2 provides summary statistics of the data.

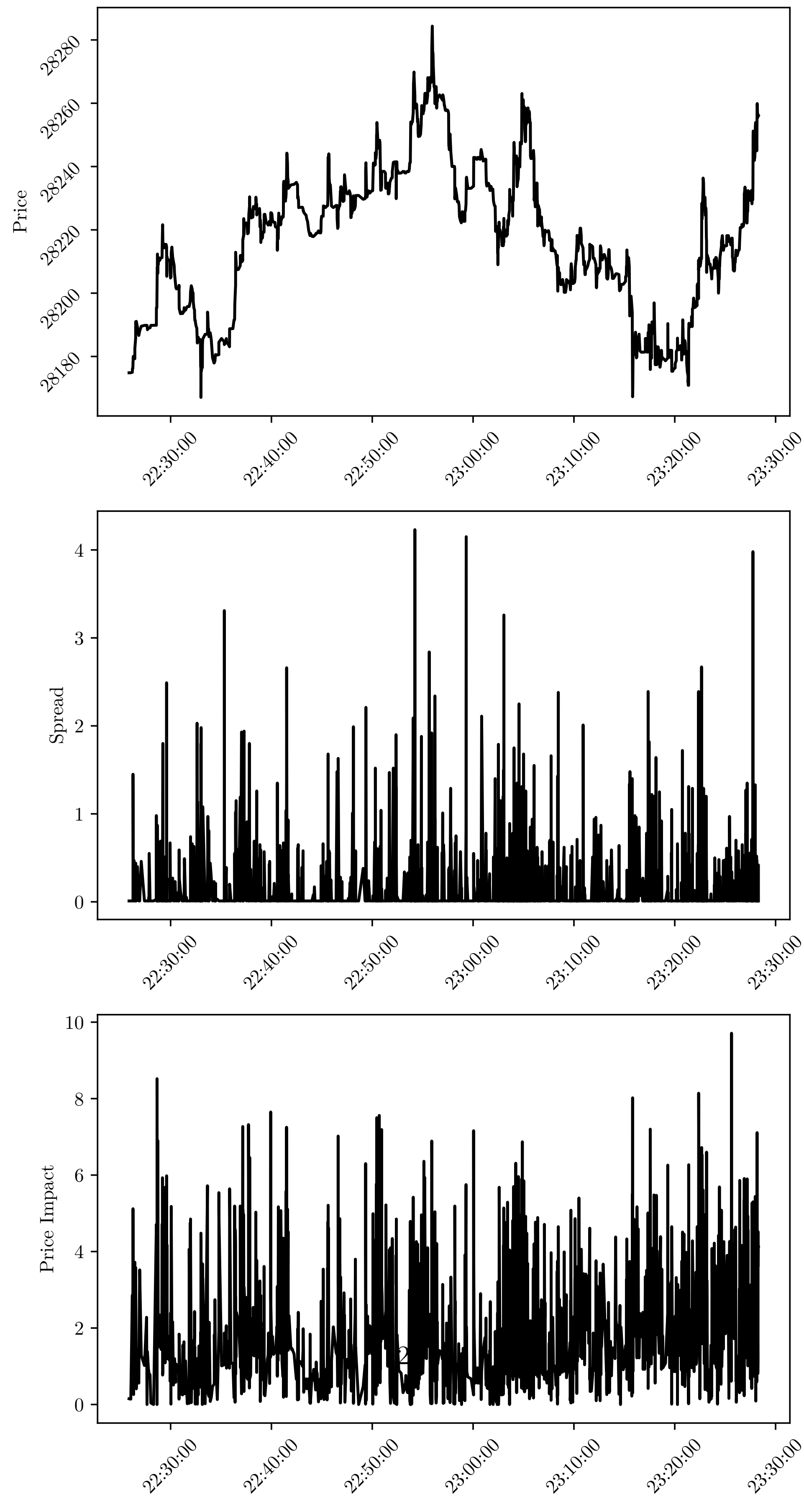
During this period, the best available ask price varied from a low of \$28,167.07 to a high of \$28,284.36, which represents a one-hour swing of about 0.38%. The mean and median price were essentially centered in that range. The spread averaged \$0.16, but over half of the regimes had a spread of \$0.01, making the distribution highly skewed. The price impact averaged a \$2.12 increase from a purchase of 2 bitcoins, but this distribution is also left-skewed, with a median price impact of \$1.81. As seen in Figure 2, the price shows a minor trend over the hour, where spread and price impact are more noisy.

3.2 Interpreting data through a steady-state lens

We begin our estimation procedure by calibrating values of ρ and θ. With the rapid pace of transactions in this exchange (with time in units of seconds), we set the rate of time preference to $\rho = 0$. Meanwhile, the average regime duration of 1.5 seconds is used to calibrate the rate of information obsolescence as $\theta = 0.66$.

We next examine across regimes for the largest order execution in a regime, which is 10 BTC. We assume that larger orders, while possible, would be deferred to a future period so as to avoid the high price impact (as in the model). This quantity $Q = 10$ corresponds to

Figure 2: Evolution of Price, Spread and Price Impact for Regime-Averaged Orderbooks



$q = 1$ in equation (16). We then convert the quantity of shares x available on the orderbook (for varying prices) to a fraction $q = x/Q$.

Next, after each orderbook update, we compute the price impact for various transaction sizes and divide them by the current spread to obtain the ratio $\frac{PI(q)}{S}$ for all q. When this data is combined with ρ and θ, equation (20) allows us to solve for $F(q)$, the distribution of the ratio of buyers to sellers.

This procedure only allows us to estimate $F(q)$ for $q \leq 1$. Our model allows for the possibility that the realized q is greater than 1, meaning that the number of shares demanded by buyers exceeds the number of shares available from sellers (at an acceptable price). These excess buyers will not be observable because they do not generate a transaction. However, we do obtain an estimate for $F(1)$, and $1 - F(1)$ indicates the fraction of realizations with excess buyers.

The next step is to use the estimates of $F(q)$ to calculate X_m and X_t in equation (12), which in turn appear in the liquidity measure L from equation (15). We note that both X_m and X_t include the expression:

$$\chi \equiv \int_1^\infty \frac{1}{q} F'(q) dq. \tag{21}$$

For an interpretation, note that $X_t = F(1) + \chi$ would the the average fraction of buyers who complete their purchase in a period, so χ is the portion of the average where buyers are in excess. Realizations of $q > 1$ are not observable, but would be censored at $q = 1$, thus χ is a free parameter. However, it necessarily must lie between 0 (if $F(q)$ were concentrated on very high qs) and $1 - F(1)$ (if $F(q)$ were concentrated on q near 1). We proceed by setting $\chi = (1 - F(1))/2$. After calculating L, we can decompose the spread into the components ΔV and L using the spread in equation (18). In so doing, our estimation of L leads to 55 out of the 2480 regimes having a negative value of L. These regimes represent situations where the assumptions of the model are violated. Specifically, in each case they occur when the there is a large gap between the lowest price and the next available price on the orderbook. When this second-best price has a small amount of liquidity, the price impact becomes very large for a small change in q. In such cases, the estimation of X_m is very large which leads to a negative value of L. We exclude these regimes from our analysis.

3.3 A correction for orderbook bias

The orderbook predicted by the model depicts the steady state distribution of available prices. At any moment, however, the observed orderbook will temporarily deviate from this steady state while some trades are transacted before being replenished by new limit orders. Indeed, market orders always cross with the most favorable limit order, so at any random moment of time, a snapshot of the ask orderbook will be biased towards higher prices, relative to the underlying steady state distribution. The observed distribution is missing the weight that has already been executed and has not yet been refilled.

To correct this bias in our estimate of the steady-state orderbook, we first take our snapshots of the orderbook at each point in time t, expressing the orderbook at price a as a cumulative distribution $F_t(a)$ by dividing the volume of limit orders at or below price a by the total volume of limit orders offered on the book. In practice, each distribution occurs on a discrete grid of prices. The observed grid of prices for each orderbook in the steady state varies since some have orders that have been executed. Let $[a^1, a^2, \ldots, a^I]$ denote the grid of all possible ask prices in this steady state. This grid is the union of all observed ask prices. Let a_{0t} denote the lowest price in the support of F_t.

Start with the lowest price on the grid, a^1. There is no (observable) missing weight for any F_t whose lowest price $a_{0t} = a^1$, so let $\hat{F}_t(a) = F_t(a)$ for all a, and then let $\bar{F}(a^1) = mean_t\{\hat{F}_t(a^1)\}$.

Proceeding across the grid of prices, for all distributions F_t whose lowest price $a_{0t} = a^i$, we then adjust for the missing weight below a^i be letting $\hat{F}_t(a) = \bar{F}(a^{i-1}) + (1 - \bar{F}(a^{i-1}))F_t(a)$, and compute their average at price a^i as $\bar{F}(a^i) = mean_t\{\hat{F}_t(a^i)\}$.

This procedure assumes that each of the observed snapshots come from the steady state orderbook, only truncated up to whichever limit orders have been executed. Averaging the snapshots is desirable because of the inherent lumpiness of actual limit orders, which can be smoothed out in this procedure. At the same time, the reweighting for missing truncated weight ensures that the resulting CDF $\bar{F}(a)$ is an unbiased estimate.[7]

3.4 Four Case Studies

To explore more carefully the empirical content of the model, we select four regimes that lasted at least 15 seconds from the sample. Characteristics for these regimes are given in table 3. From equation (19) we see that price impact can be decomposed into the terms $(\rho + \theta)\frac{\Delta V}{L}$ and $\frac{F(q)}{1-F(q)}$. Given the characteristics of the sample shown in the last section, we calibrate the model with the values $Q = 10$ and $\rho + \theta = 0.67$. These values remain constant throughout the remainder of the paper. The value of Q is chosen to be large enough that it is unlikely to be reached in the sample, but small enough that it is plausible that it could be reached. The value of $\rho + \theta$ is chosen to be consistent with the average duration of the regimes in the sample. Since we don't observe 10 BTC being executed per second in our sample, we calibrate the model with $\chi = 0$.

These four regimes have similar durations (between 17 and 22 seconds), but differ in the average spread over the period as well as the number of executions observed and the quantity of BTC transacted per second. The model calculated probability of 2 BTC or less being transacted per second (which is derived from the equilibrium conditions and the observed orderbook, as opposed to observed executions) varies from between 68.4% and

[7]This process is similar to Kaplan-Meier (KM) adjustment for left-truncated data. The key difference is that KM would take each limit order as the unit of observation, rather than each orderbook. If the number of limit orders in each book were the same, the processes would be identical.

Table 3: Descriptive statistics of selected regimes

	Avg. Spread	Duration	Number of executions	BTC executed per second	Price Impact ($q = 0.2$)	ΔV	L	$F(q = 0.2)$
1	0.46	17.90	66	0.496	1.29	6.16	26.77	0.830
2	0.38	17.66	26	0.016	0.47	2.35	12.35	0.684
3	0.03	18.52	66	0.129	0.13	3.56	237.29	0.883
4	0.01	21.31	32	0.015	2.03	4.91	982.39	0.997

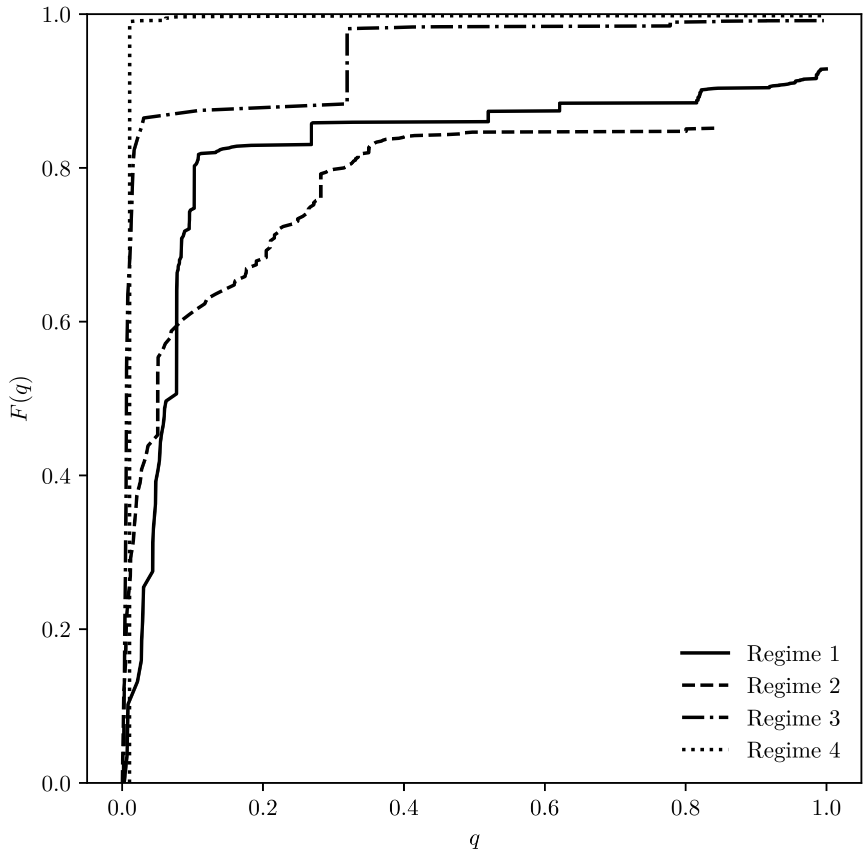
99.7%.

Consider the two regimes that have the highest average spreads (1 and 2). These two regimes have approximately the same duration, while regime 1 is more traditionally illiquid in that it has a large spread (\$0.46) and a large price impact (\$1.29). While regime 2 has nearly as large an average spread, the price impact is smaller than in regime 1 (\$0.47). The model attributes this difference to regime 1's higher ratio of valuation component (ΔV) to liquidity component L.

In contrast, regimes 3 and 4 have smaller spreads than regimes 1 or 2, leading to relatively large market liquidity measures L. Interestingly, regime 4 has the largest price impact, resulting in a large $F(0.2)$ estimate. This suggests that sellers were expecting a low flow of executions. Indeed, the difference between regimes 3 and 4 was largely a function of changes in the liquidity measure L, which is derived from changes in the shape of the orderbook (i.e. price impact), as opposed to changes in the spread. In contrast, regimes 1 and 2 differ mainly because of the difference in valuations of the market participants. Figure 3 shows the full estimated distribution of order flows, $F(q)$, for each of these regimes.

These examples demonstrate that the components that determine liquidity, ΔV, L and $F(\cdot)$ can vary across steady-state regimes in an asset market and a greater understanding of liquidity can be obtained by understanding each of these components.

Figure 3: Estimated distribution of order flows, $F(q)$, for selected regimes



3.5 The contribution of valuation differences and market frictions in explaining spread and price impact

This subsection presents distributional information about the contribution of valuation differences and the liquidity measure L in determining spreads across all steady-state regimes.

Figures 4 and 5 show the variation in the valuation component and the market friction component, respectively, across the regimes described in the previous subsection for all regimes.

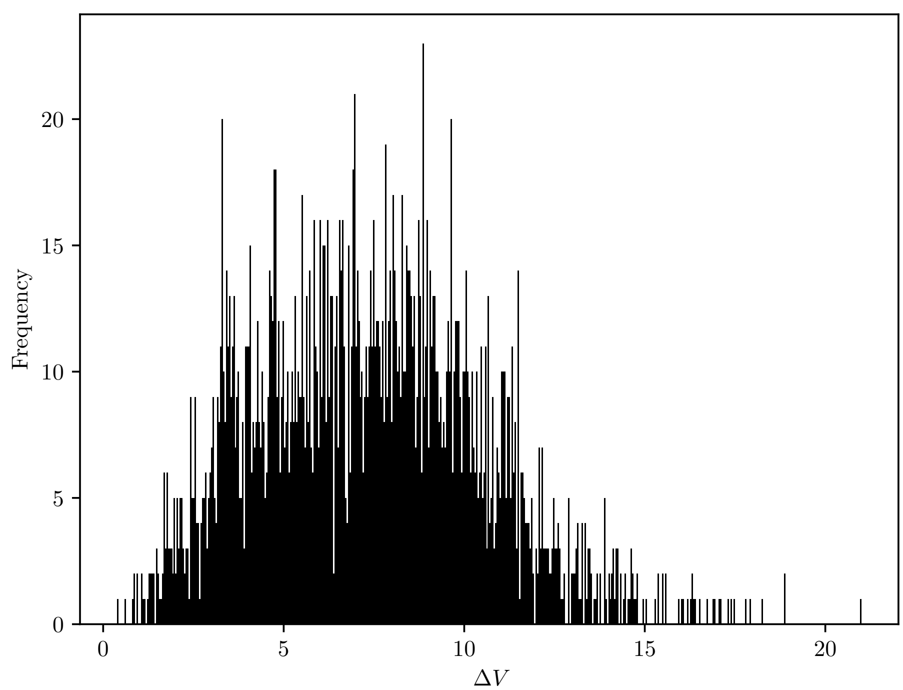


Figure 4: Distribution of ΔV across steady-state regimes

Figures 6 and 7 show the variation in the valuation component and the market friction component, respectively, over time for $\chi = 0$.

Given the distinction between the valuation component of spread and the liquidity com-

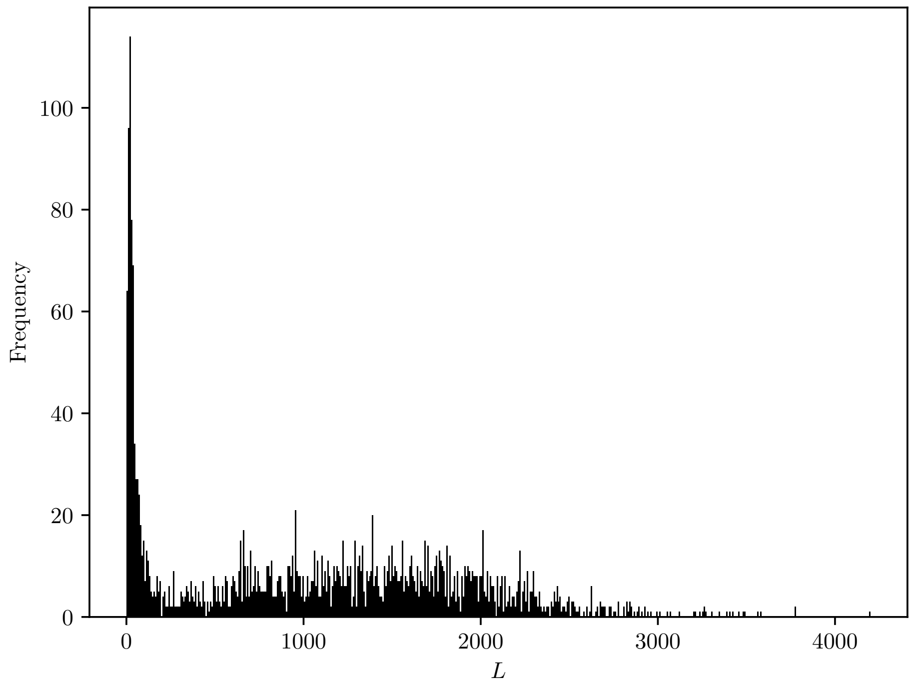


Figure 5: Distribution of L across steady-state regimes

ponent of the spread, we now turn to the question of understanding the extent to which variation in the spread is driven by changes in ΔV or changes in L. Figure 8 plots a histogram of the valuation component ΔV for regimes where the spread is \$0.02 or less (in green), vs regimes where the observed spread is greater than \$0.02 (in red). These distributions appear similar. A Kolmogorov-Smirnov test of the null hypothesis that the two distributions are the same cannot be rejected at the 5% level. This suggests that the valuation component of the spread is not driving the variation in the spread.

Figure 9 plots a histogram of the liquidity component L for regimes where the spread is \$0.02 or less (in green), vs regimes where the observed spread is greater than \$0.02 (in red). The stark difference in these distibutions suggests that changes in the L drive the variation in the spread. The figure shows that when the spread is greater than \$0.02, the

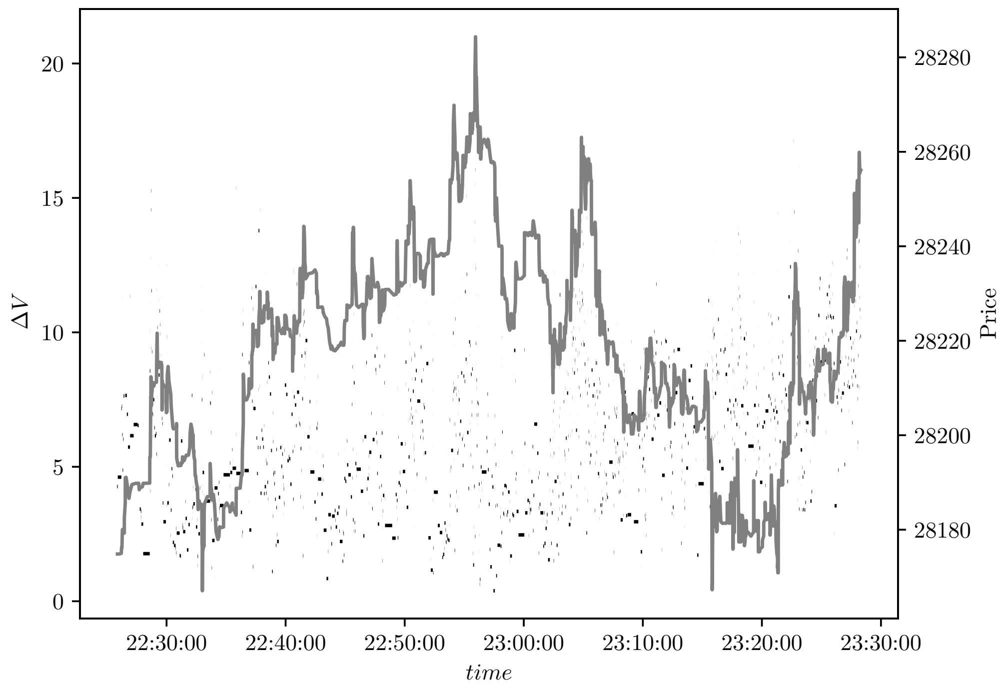


Figure 6: ΔV and price over time

market friction component is always small. However, when the spread is \$0.02 or smaller, the liquidity component L varies substantially.

Since $S = (1 + \rho + \theta)\Delta V/L$, and L is calculated directly, while ΔV is the residual of S not explained by L, ρ and θ, we seek to understand the statistical relationship between S, ΔV and L. The model suggests that $\log S$, $\log \Delta V$ and $\log L$ should be linearly related. The correlation matrix between these three variables is given in table 4. The correlations corroborate the story presented in figures 8 and 9. The correlation between $\log S$ and $\log \Delta V$ is 0.028 while the correlation between $\log L$ and $\log S$ is -0.959. This demonstrates that the information encoded from the orderbook through L accounts for a large fraction of the variation in the spread. A regression of the log spread on the log liquidity component (plus a constant) yields a significant coefficient of -0.902 and an R^2 of 0.892. Of interest, a regression

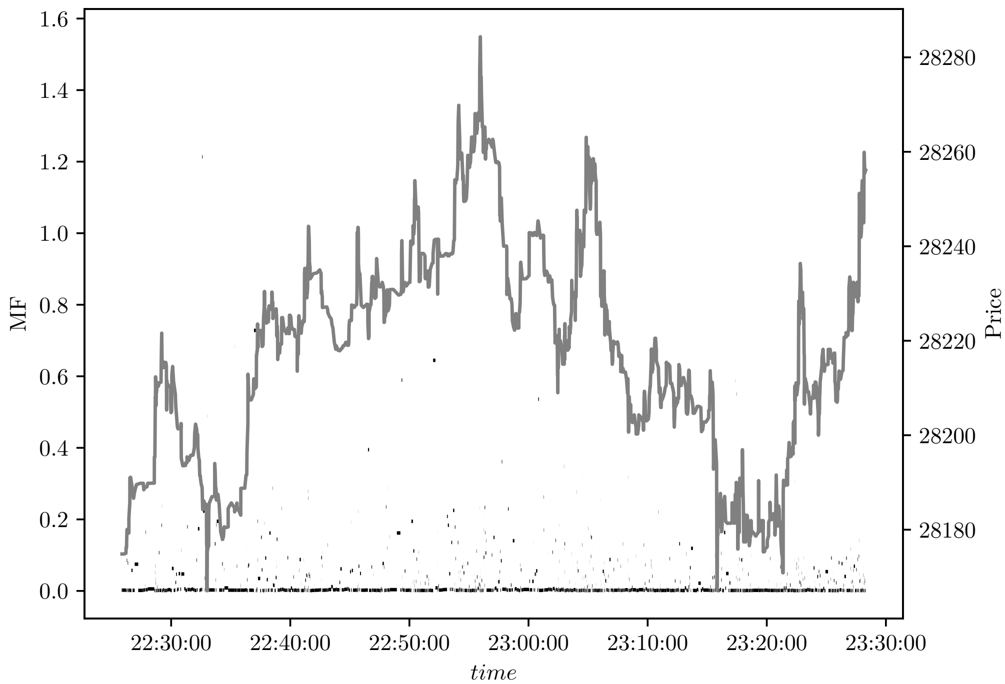


Figure 7: L and price over time

of the log price impact (at $q = 0.2$) on $\log S$ has an R^2 of 0.0015.

3.6 Consistency of model predictions with data

One way to verify external validity of the model is to use executions data for the period in question and correlate it with the model-implied distribution of executions found in F. As is often the case in models that involve investor expections, checks on the consistency of the model with observed data are a joint test of many model assumptions, including the correctness of individuals' expections. If individuals' expectations are correct, then in regimes where $F(q)$ is high for a given q, we would expect to see a low number of executions, since $F(q)$ gives the implied probability that the fraction of the orderbook executed will be less than q.

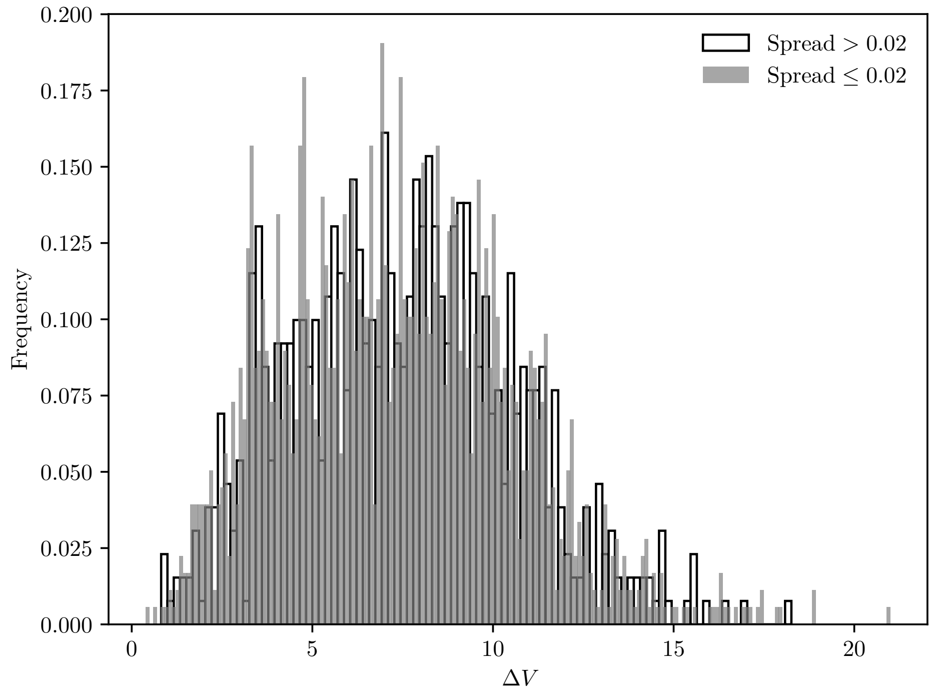


Figure 8: ΔV for spreads less than \$0.02 (green) and greater than \$0.02 (red)

A regression of $F(0.2)$ on BTC executed per second yields a coefficient of -0.259 with a t-statistic of 4.129 and an R^2 of 0.079 when the sample includes all regimes that last at least 5 seconds. This result is consistent with the model's predictions.

4 Welfare

We now consider the gains from trade in an orderbook market. First, we establish two extremes with no trade and frictionless trade. Autarky would leave the asset with sellers indefinitely, generating a total welfare of $\frac{d_A}{\rho}$ in utils for the seller. If there were no frictions from the timing of trades, then each seller who enters the market would immediately find a buyer, generating a total welfare of $\frac{d_B}{\rho} - f_t - f_m$. The gains from this frictionless trade

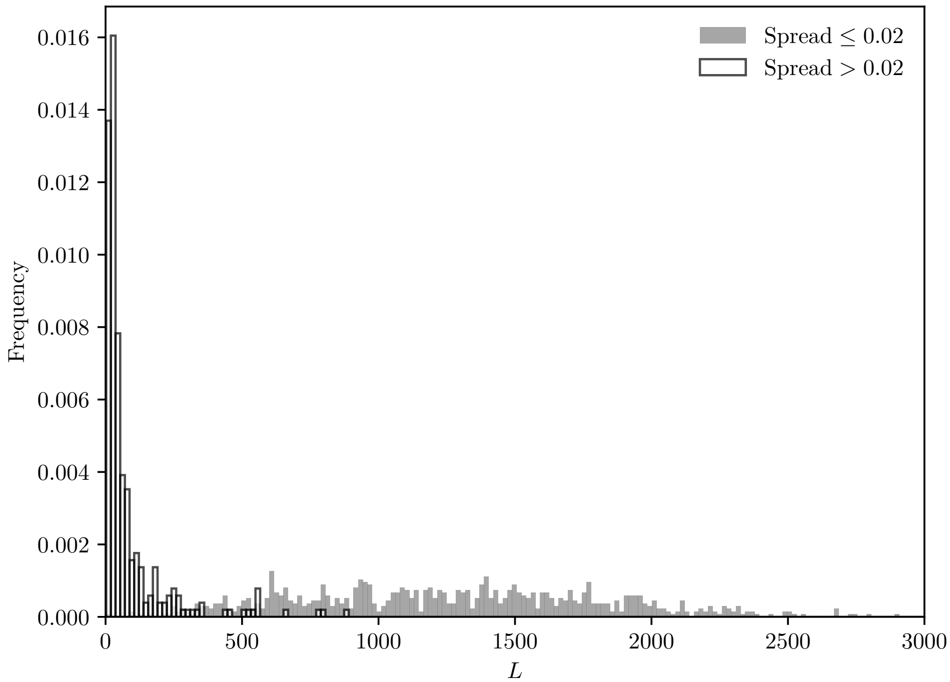


Figure 9: L for spreads less than \$0.02 (green) and greater than \$0.02 (red). Limiting the L (horizontal) axis to 3000 removes 24 of the 2479 regimes.

would then be their difference, which we have already defined as ΔV.

We compare this frictionless benchmark to the gains in an orderbook market. Each buyer who enters the market receives expected utility W_B and each seller receives V_S. We add these to obtain:

$$\begin{aligned} TW &= \frac{d_B}{\rho} - f_t - a_1 + \frac{a_0 + f_m + d_A}{1+\rho+\theta} \\ &= \frac{d_B}{\rho} - f_t - f_m - \frac{\Delta V}{\frac{\rho+\theta+X_t}{1-F(1)} + 1 - X_m - X_t}\left(\frac{\rho+\theta}{1-F(1)}\right) \end{aligned}$$

We can then consider what fraction of the potential gains from trade are realized in the h

	$\log S$	$\log \Delta V$	$\log L$
$\log S$	1.000000	0.028488	-0.958723
$\log \Delta V$	0.028488	1.000000	0.256913
$\log L$	-0.958723	0.256913	1.000000

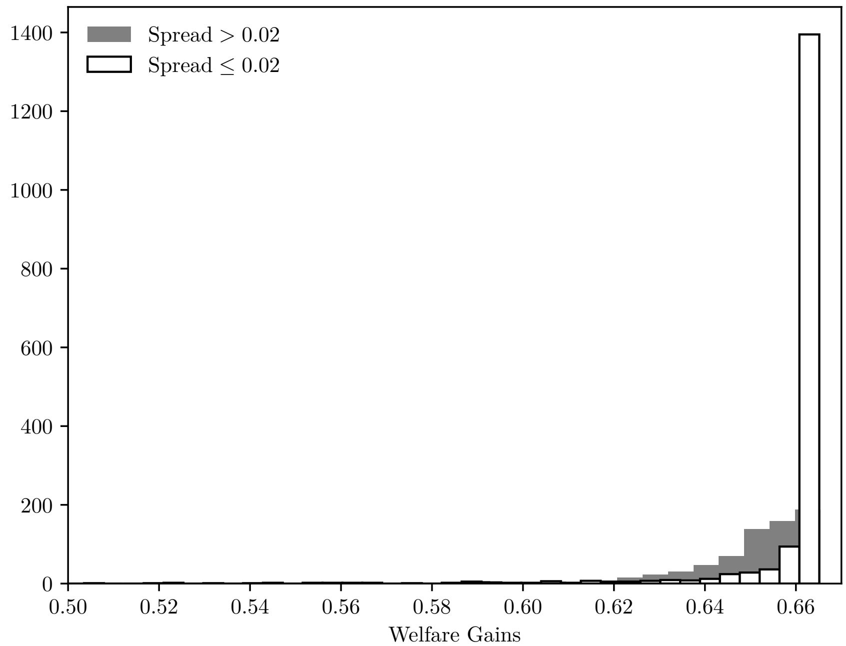
Table 4: Correlations between log spread and its components

market, which would be:

$$\begin{aligned}
\%Gain &= \frac{TW - \frac{d_A}{\rho}}{\Delta V} \\
&= \frac{\Delta V - \frac{\Delta V}{\frac{\rho+\theta+X_t}{1-F(1)}+1-X_m-X_t}\left(\frac{\rho+\theta}{1-F(1)}\right)}{\Delta V} \\
&= \frac{\frac{X_t}{1-F(1)} + 1 - X_m - X_t}{\frac{\rho+\theta+X_t}{1-F(1)} + 1 - X_m - X_t} \\
&= \frac{L - \frac{\rho+\theta}{1-F(1)}}{L}
\end{aligned} \tag{22}$$

The fraction of realized gains are entirely driven by the size of the liquidity component relative to the effective discount rate $(\rho+\theta)/(1-F(1))$; the known difference in buyer and seller valuations only affects the total level of gains. This fraction lies between 0 and 1. If ρ and θ are both small, market participants are patient and their information is long-lived, so they are unconcerned by the delayed process of transacting. Thus, most of the potential gains are realized. The remaining terms measure how frequently the realized number of buyers and sellers are unequal. Imbalance in either direction is inefficient because more time is required to transfer the asset from low- to high-valuation participants. Extra sellers will bring down average welfare because they receive low utility while awaiting buyers, and excess buyers will bring down the average welfare because they receive no utility until they transact. Figure 10 shows the distribution of percent gains across regimes, divided by spreads greater

Figure 10: Distribution across steady-state regimes of the realized fraction of potential welfare gains. Limiting the x axis to 0.6 and above removes 20 of the 2479 observations.



than or less than or equal to 0.02.[8]

To illustrate this, consider two distributions of participants, F and $\hat{F}$, where $F(q) = \hat{F}(q)$ for all $q \leq 1$, and $F(q) < \hat{F}(q)$ for all $q > 1$. That is, F provides an equivalent distribution as $\hat{F}$ when buyers are the short side of the market, but when sellers are the short side of the market, F is more likely to have larger numbers of un-served buyers. With rearrangement of the components of X_t and X_m, we get:

$$\begin{aligned} \frac{X_t}{1-F(1)} + 1 - X_m - X_t \quad &= \quad 1 - \frac{F(1)^2}{1-F(1)} - \int_0^1 \frac{F'(q)}{q} \left(\int_0^q \frac{F(s)}{1-F(s)} ds \right) dq \\ &\quad + \left(\int_1^\infty \frac{F'(q)}{q} dq \right) \left(\frac{F(1)}{1-F(1)} - \int_0^1 \frac{F(s)}{1-F(s)} ds \right) \end{aligned}$$

[8]Given the discreteness of the data and our continuous approximation of it, there are 13 regimes for which the numerator in equation (22) is negative. We leave these out of the figure as well as trimming the x-axis to allow for a better understanding of where the heaviest weight of the distribution lies. This trimming removes the 13 values that are negative as well as 40 other observations that have a percent gain of less than 0.5.

Note that the first line will be the same under F or $\hat{F}$. In the second line, the first integral will be larger under $\hat{F}$ as it places more weight where $1/q$ is bigger. The last parenthetical element is positive because $F(1) > F(s)$ for all $s < 1$, so $\frac{F(1)}{1-F(1)} > \frac{F(s)}{1-F(s)}$.

A similar comparison can be made if $F(q) > \hat{F}(q)$ for all $q < 1$, and $F(q) = \hat{F}(q)$ for all $q \geq 1$. That is, when buyers are the short side of the market, F is more likely to have larger numbers of un-served sellers. In the second line, $\frac{F(s)}{1-F(s)} > \frac{\hat{F}(s)}{1-\hat{F}(s)}$, so the parenthetical term is larger under $\hat{F}$. In the first line, the second integral is smaller for the same reason for each q, and furthermore $\hat{F}'$ will place more weight where there is a smaller $1/q$. Thus the double integral is smaller, and with the negative, the whole term is larger under $\hat{F}$.

Thus, as the distribution F places more weight near $q = 1$ (on either side), the market friction is lessened and more gains from trade are realized.

5 Conclusion

This paper builds a model of limit order books that characterizes both spreads and price impact in general equilibrium. Trade is motivated because buyers and sellers have differing opinions of the fundamental asset value, but all buyers share the same view of that value, and similarly within all sellers. Because trades are organized in an orderbook, sellers are more likely to transact when placing a lower limit order, but get more money from the transaction at a higher limit order. This generates dispersion in transaction prices, not because of disagreement in asset value, but due to random variation in the flow of buyers, where high flows eat further into the orderbook and lead to higher prices.

This leads to novel predictions that suggest it is informative to study the spread and price impact — both common measures of market frictions — jointly so as to interpret whether changes in valuation or in order flows are driving market prices. We apply the model to data from the Coinbase BTC/USD exchange, finding plausible time periods (regimes) during which the market appears to be in steady state, and use the model to interpret differences among them.

References

Back, K. & Baruch, S. (2004), 'Information in securities markets: Kyle meets glosten and milgrom', *Econometrica* **72**(2), 433–465.

Butler, A. W., Grullon, G. & Weston, J. P. (2005), 'Stock market liquidity and the cost of issuing equity', *Journal of Financial and Quantitative Analysis* **40**(2), 331–348.

Colliard, J.-E. & Foucault, T. (2012), 'Trading Fees and Efficiency in Limit Order Markets', *The Review of Financial Studies* **25**(11), 3389–3421.
URL: *https://doi.org/10.1093/rfs/hhs089*

Cui, W. & Radde, S. (2016), 'Money and asset liquidity in frictional capital markets', *American Economic Review* **106**(5), 496–502.

Foucault, T., Kadan, O. & Kandel, E. (2005), 'Limit order book as a market for liquidity', *The review of financial studies* **18**(4), 1171–1217.

Glosten, L. R. & Milgrom, P. R. (1985), 'Bid, ask and transaction prices in a specialist market with heterogeneously informed traders', *Journal of financial economics* **14**(1), 71–100.

Goettler, R. L., Parlour, C. A. & Rajan, U. (2005), 'Equilibrium in a dynamic limit order market', *The Journal of Finance* **60**(5), 2149–2192.

Kerr, J., Sadka, G. & Sadka, R. (2020), 'Illiquidity and price informativeness', *Management Science* **66**(1), 334–351.

Kyle, A. S. (1985), 'Continuous auctions and insider trading', *Econometrica: Journal of the Econometric Society* pp. 1315–1335.

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

Pástor, L. & Stambaugh, R. F. (2003), 'Liquidity risk and expected stock returns', *Journal of Political economy* **111**(3), 642–685.

Vayanos, D. & Wang, T. (2007), 'Search and endogenous concentration of liquidity in asset markets', *Journal of Economic Theory* **136**(1), 66–104.