ESTIMATING GAINS FROM TRADE IN FINANCIAL MARKETS

NIHAD ALIYEV [a,*], ROBERT GAUDIOSI [a,*], AND TALIS J. PUTNINS [b,c,d,*]

[a] Department of Applied Finance, Macquarie University, Sydney

[b] University of Technology Sydney

[c] Stockholm School of Economics in Riga

[d] Digital Finance Cooperative Research Centre

Abstract. We develop a method to estimate gains from trade in financial markets. The method builds on the idea that as markets become more liquid, more traders—especially those with private values near the midprice—begin to participate in trading, increasing total gains. Our method is simple, requires minimal assumptions, and can be used to evaluate the potential impact of market reforms on total gains from trade. Applying the method to Decimalisation, Autoquote, and the Tick Size Pilot (TSP), we find that Decimalisation increased annual trader gains by $39–78 million on NYSE and $1.2–2.3 billion on NASDAQ; Autoquote generated $90–180 million in annual trader gains on NYSE; and the TSP reduced trader gains by $53–106 million annually. Gains are most sensitive to trading costs in small-cap, volatile, and algorithmically traded stocks. The method offers policymakers and researchers a tool to evaluate the economic importance of market reforms.

JEL Classification: D47, D61, G10, I38

Keywords: Gains from trade, market quality, liquidity, volume

---

Date: September 29, 2025.

(*) Nihad Aliyev (nihad.aliyev@mq.edu.au), Robert Gaudiosi (robertisaac.gaudiosi@hdr.mq.edu.au), Talis J. Putnins (talis.putnins@uts.edu.au).

1. Introduction

Suppose a proposed market design or regulatory change costs $100 million per year in compliance and monitoring, and reduces bid-ask spreads by 10 basis points. Is it socially beneficial? While the premise is straightforward, the academic literature lacks tools to reliably answer this question. Markets improve welfare by enabling mutually beneficial trades, yet we lack empirical methods to quantify how much better off participants become i.e., the gains from trade. This gap becomes critical when evaluating efficiency-enhancing reforms: if lower trading costs encourage participation and increase gains from trade, we must be able to measure those benefits to justify policy costs.

This paper addresses that gap by developing a simple and practical method to estimate gains from trade in financial markets. These gains are the direct analogue of economic surplus in goods markets—a concept that lies at the heart of welfare economics. From Harberger's (1966) deadweight-loss triangles to Hausman's (1981) empirical surplus measures, surplus has long provided a benchmark for evaluating policy and market efficiency. By framing financial market activity in terms of surplus, we situate our measure of gains from trade within this well-established tradition of welfare analysis. Unlike structural models that require granular data and strong distributional assumptions (e.g., Hollifield et al., 2006; Goettler, Parlour & Rajan, 2005), our approach relies only on readily available data—bid-ask spreads and trading volumes—and is computationally feasible at scale.

Our approach is grounded in a fundamental property of market exchange: traders have heterogeneous private values, and only those whose value differentials exceed the cost of trading will participate. Private values arise because traders differ in their motives for transacting. For instance, some trade to diversify portfolios, hedge risks, or transfer resources across states of the world or across time—each of which generates benefits that make an asset worth more (or less) to them than the prevailing market price. These differences in valuation are precisely what create the welfare gains from trade: a buyer's willingness to pay above the midpoint reflects the benefit they derive from acquiring the asset, while a seller's willingness to accept below the midpoint reflects the benefit of reallocating resources. While some gains from trade are realized through existing trades, others remain unrealized due to non-participation. Reductions in trading costs increase market participation (volume), thereby increasing realized gains from trade. By estimating the relationship between trading volume and bid-ask spreads, we can recover implied private values and compute marginal gains from trade.

An additional benefit of our method is that it decomposes gains from trade into two components: the surplus accruing to traders and the revenue accruing to market makers. This distinction clarifies how policy changes affect different market participants. Trader surplus is straightforward: because traders pay the bid-ask spread, they always gain from tighter spreads, which expand the range of private value differentials that make trade worthwhile. Market makers, by contrast, earn their revenue from the spread. They benefit from wider spreads when trading volume is relatively inelastic, but narrower spreads can increase their revenue if the accompanying growth in volume is sufficiently elastic to offset lower per-trade margins.

The analogy to standard welfare analysis is that market maker revenue corresponds to tax revenue—the transfer between traders that accrues to an intermediary. However, unlike taxes, the bid-ask spread compensates market makers for both fixed costs (e.g., exchange fees, brokerage costs) and variable costs such as adverse selection and inventory risk (e.g., Glosten & Milgrom, 1985; Goldstein & Hotchkiss, 2020; Kyle, 1985; Liu & Wang, 2016). For this reason, we interpret "market maker gains" as market maker revenue, which may largely reflect the costs of doing business rather than pure profit. In the analysis that follows, we therefore use the term market maker revenue, with increases (resp. decreases) in revenue understood as potentially reflecting decreases (resp. increases) in underlying costs.

We apply our approach to three major market design changes: Decimalisation, Autoquote, and the Tick Size Pilot (TSP) Q Rule. We estimate that annual trader gains from Decimalisation were \$39–78 million on NYSE and \$1.157–2.314 billion on NASDAQ. Market maker revenue, by contrast, declined by \$5 million on NYSE and \$1.191 billion on NASDAQ. Gains depend on the proportion of trades intermediated by market makers, with the lower bound corresponding to a pure dealer market and the upper bound to a limit order market with no pure market makers.[1] Trader gains from NYSE Autoquote ranged from \$90 to \$180 million annually, with market maker revenue dropping up to \$72 million. In contrast, the TSP Q Rule imposed annual trader losses of \$53–106 million, while market makers revenue increased up to \$66 million on NYSE and \$78 million on NASDAQ. The opposing signs between trader gains and market maker revenues highlight the relative inelasticity of volume to bid-ask spreads on these exchanges.

[1] Intermediate values depend linearly on the proportion of market makers from the lower bound (100% market makers) to the upper bound (0% market makers)

In addition to enabling *ex-post* evaluation, the method can also be used to assess the welfare effects of a market design change *ex-ante*, provided policymakers can anticipate its impact on bid-ask spreads. Finally, we explore extensions: estimating total (not just marginal) gains from trade, and examining which stock characteristics predict the greatest welfare gains. We find that gains are most sensitive to trading costs in small-cap, volatile, and algorithmically traded stocks.

Beyond policy evaluation, our framework helps quantify the economic significance of empirical microstructure results. Many studies document changes in spreads from market reforms-but without a welfare benchmark, it is difficult to assess economic significance.[2] Our approach fills this gap, enabling comparisons across asset classes and helping to prioritize regulatory and research efforts.

1.1. Related Literature. This paper builds on a limited set of studies that estimate gains from trade in limit order markets. Hollifield et al. (2006) and Goettler, Parlour, and Rajan (2005) use rich structural models to infer trader private values from order submissions, cancellations, and executions. These methods link gains from trade to optimal order submission strategies and require strong assumptions (e.g., on distributions of private values and event timing) and detailed order book data. As a result, they are difficult to apply at scale. In contrast, our approach links gains from trade to observed variation in trading volume and bid-ask spreads rather than inferred strategic behaviour. This makes our method model-free, computationally light, and broadly applicable using standard market data.

This study is also related to experimental studies of gains from trade in double auction markets (e.g., Cason & Friedman, 1996; Kagel, 2004). In such studies, private values are assigned exogenously to traders, so that private values are observable and gains from trade can be computed. Our contribution differs by estimating gains empirically in real markets where private values are latent.

Finally, our paper aligns with work in other domains that estimate welfare using demand curves. For instance, Einav, Finkelstein, and Cullen (2010) and DeFusco, Tang, and Yannelis (2022) exploit price variation in health insurance and credit markets, respectively, to estimate demand and cost curves and quantify welfare losses from adverse selection. In those contexts, welfare losses arise from marginal costs that increase with adverse selection. In securities markets, similar inefficiencies arise from asymmetric information, reflected in the bid-ask spread (Glosten & Putnins, 2016).

[2]Notable studies of this type include Bessembinder, (2003); Brogaard et al. (2015); Hendershott, Jones & Menkveld (2011); Chung, Lee & Rosch (2020); Foley et al. (2022); Riordan & Storkenmaier (2012).

2. Conceptual framework

2.1. Gains from trade. Consider the market for a tradeable security with T traders arriving sequentially over a finite horizon. Each trader has a private valuation for the security of α_t. Trade occurs between a buyer and seller when $\alpha_t - p_t^B > 0$ (buyer) or $p_t^{Sell} - \alpha_t^S < 0$ (seller) where p_t^B is the price available to buy and p_t^S is the price available to sell. If the potential gain from trade is negative than a trader abstains from trade.

2.1.1. *Limit-order market.* In a standard limit order book without designated market makers, any trader can supply liquidity. We distinguish between two types of liquidity providers: *patient traders* and *market makers* (analogous to dealers in a dealer market). Patient traders possess private values α_t and trade under the same conditions as any other traders: they buy if $\alpha_t - Ask_t > 0$ or sell if $Bid_t - \alpha_t > 0$. Market makers, by contrast, have private values equal to the midquote and earn the half-spread on each transaction.[3]

Let MM_t be a dummy variable equal to 1 if the passive liquidity provider in transaction t is a patient trader, and 0 if it is a market maker. The total gains from trade accruing to traders are:

$$TraderGains_T = \underbrace{\sum_{t=1}^{T} \left(\tfrac{1}{2} + MM_t\right) \left(\alpha_t^B - Ask_t\right) \delta_t^B}_{\text{Aggressive Buys}} + \underbrace{\sum_{t=1}^{T} \left(\tfrac{1}{2} + MM_t\right) \left(Bid_t - \alpha_t^S\right) \delta_t^S}_{\text{Aggressive Sells}}. \quad (1)$$

The total gain is the sum of buyer and seller gains. Conditional on participation, the aggressive side of the trade always earns the difference between the trade price and their private value. The scaling factor $\left(\frac{1}{2} + MM_t\right)$ captures that the passive side of the trade shares in these gains if it is a patient trader ($MM_t = 1$).

Market makers earn revenue from the bid-ask spread. Their total revenue is:

$$MMRevenue_T = \underbrace{\sum_{t=1}^{T}(1 - MM_t) \times \delta_t^B \times \frac{Ask_t - Bid_t}{2}}_{\text{Aggressive Buys}} + \underbrace{\sum_{t=1}^{T}(1 - MM_t) \times \delta_t^S \times \frac{Ask_t - Bid_t}{2}}_{\text{Aggressive Sells}}. \quad (2)$$

As discussed, this revenue may reflect compensation for the costs of doing business—such as exchange fees, adverse selection, and inventory risk—rather than pure economic surplus.

[3] This formulation is consistent with canonical microstructure models. For example, in Glosten & Milgrom (1985), the market maker's unconditional belief about the asset value is the midquote.

2.1.2. *Dealer market.* In a dealer market, one side of every transaction is always an intermediary. This is a special case of the limit order market with $MM_t = 0$ for all t.

In this framework, Eq. (1) provides the range of possible trader gains. When $MM_t = 0$ for all t (a pure dealer market), it represents the lower bound. When $MM_t = 1$ for all t (a pure limit-order market with only patient traders supplying liquidity), it represents the upper bound. Conversely, Eq. (2) characterizes market maker revenue. The lower bound is zero when $MM_t = 1$ for all t, while the upper bound is reached when $MM_t = 0$ for all t.

2.2. Graphical representation. In this section, we present a supply–demand framework to illustrate gains from trade and the welfare cost of illiquidity. The distribution of trader private values (α_i) is reflected in the shape of the demand and supply curves. We assume that buyers and sellers draw from identical private value distributions, so the supply curve is the mirror image of the demand curve. Under this assumption, we do not differentiate between gains accruing to buyers and those accruing to sellers.

Let the bid-ask spread be S and dollar volume be $V(S)$. Consistent with the previous analysis, define the proportion of trades intermediated by market makers as: $MM = \frac{1}{T}\sum_{t=1}^{T} MM_t$. Trader gains from trade are:

$$TraderGains = (1 + MM) \times \left(\int_0^{V(S)} D(V(S), S)\, \partial V - Ask \cdot V(S) \right). \tag{3}$$

On the other hand, market makers earn half the bid-ask spread per trade, shown in Figure 1A as half the rectangle (M), scaled by $(1 - MM)$ to reflect Eq. (2). In standard welfare analysis, the entire rectangle (M) corresponds to tax revenue.[4] In standard applications, both buyers and sellers pay tax, while in our setting a trade is either intermediated by a market maker (who earns half the spread from one side) or directly matched between a buyer and a seller (no market maker intermediation). Thus, market maker revenue corresponds to a partial transfer by one counterparty (half of M). Trader welfare corresponds to the shaded triangles in Figure 1A: when a market maker is on one side of the transaction, welfare is the upper or lower triangle (T), and when trades are direct, welfare is the sum of upper and lower triangles.

[4]See Harberger (1966).

The deadweight loss due to the bid-ask spread is the unrealized triangle (DL) scaled by MM (or equivalently, half the triangle scaled by $1 + MM$), plus half the rectangle (M) scaled by $1 - MM$:

$$DeadweightLoss = (1 + MM) \times \left(\int_{V(S)}^{V_{max}} D(V(S), S)\, \partial V - P^* \cdot (V_{max} - V(S)) \right) \tag{4}$$
$$+ MM \times (V(S) \cdot (Ask - Bid))$$

This represents the welfare that would be realized if the bid-ask spread were eliminated. Maximum gains from trade occur when the bid-ask spread is zero, eliminating deadweight loss. Figure 1B shows the incremental gains from trade when the spread narrows, and Figure 1C illustrates the implied relationship between volume and the spread. The figures assume an equal number of buys and sells by non-market-maker traders. This is trivial in a pure limit-order book (no market makers), but is also reasonable with market makers present, since dealers actively manage inventories close to zero to avoid price risk (e.g., Liu & Wang, 2016).

[Figure 1 here]

For market makers, the aggregate change in revenue due to a market design change that alters the spread is:

$$\Delta MMRevenue = \frac{1}{2}(1 - MM) \times \left(S_{post}\,(V_{post} - V_{pre}) - (S_{pre} - S_{post})\,V_{pre} \right). \tag{5}$$

Two opposing effects are at play: (i) a tighter spread reduces per-trade revenue, but (ii) the associated increase in volume raises total revenue.

When the volume–spread relationship is linear (as under uniform private value distributions), trader incremental gains from such a change simplify to:

$$\Delta TraderGains = (1 + MM) \left(\frac{1}{2}\,(S_{pre} - S_{post})\,V_{pre} + \frac{1}{4}\,|(S_{pre} - S_{post})|\,(V_{post} - V_{pre}) \right). \tag{6}$$

In the empirical analysis, we exploit this simplified expression (Eq. 6) by approximating the demand curve with small linear segments.

The framework also extends to a pure dealer market ($MM = 0$). Importantly, the deadweight loss due to the spread is always larger in a limit-order book market than in a dealer market: specifically, Eq. (4) is strictly larger for $0 < MM \leq 1$ than for $MM = 0$.

3. Data and descriptive stats

3.1. Data. We source 5-minute frequency data from LSEG Tick History and daily data from Centre for Research in Security Prices (CRSP) for the universe of stocks listed on the NASDAQ and NYSE from 1 January 1996 to 31 December 2022. The LSEG data contains data for the best bid and ask at the end of each 5-minute window and traded volume (in shares) during the corresponding period. We reduce the dataset by limiting our analysis from 9:40AM to 3:50PM on each trading day, excluding the opening and closing 10 minutes. From this data, we compute dollar volume, relative bid-ask spreads, and drop negative spread observations.

Prior to analysis, a 0.1% winsorization is applied to the tails of relative bid-ask spreads and volume for each stock separately. Stocks that do not trade for more than one-quarter of days in each analysis are dropped from the sample. This non-conservative data cleaning process leaves us with a well represented sample of stocks on the NASDAQ and NYSE. For example, over our full sample period we retain 10487 stocks on the NASDAQ and 7196 stocks on the NYSE. We acknowledge that a significant portion of these stocks are relatively illiquid, small cap stocks. However, for this study we are interested in total gains from trade in the marketplace which also includes these stocks.

3.2. Description of events. We consider three major market design implementations that are shown in the literature to have impacted market liquidity (Bessembinder, 2003; Chung, Lee & Rosch, 2020; Hendershott, Jones & Menkveld, 2011) to estimate the impact on gains from trade. These events are described below.

3.2.1. *Decimalisation.* Both NYSE and NASDAQ transferred from a system of fractional pricing to decimal pricing in the early 2000's. The switch to decimal pricing coincided with a reduction in the minimum tick size in both markets to one cent, enabled by the transition. The move to decimalization had an impact on liquidity through a narrowing of bid-ask spreads, particularly pronounced in larger capitalisation stocks (Bessembinder, 2003; Chung, Van Ness & Van Ness 2004). Decimalisation was phased in from 28 August 2000 to 29 January 2001 on the NYSE and on the NASDAQ from March 12 2001 to April 9 2001. For our analysis, we use 1-week before the implementation of Decimalisation on NYSE for the pre-event window and 1-week after the implementation of decimsaliation on NASDAQ for the post-event window.

3.2.2. *Autoquote.* NYSE introduced Autoquote in 2003. Prior to the implementation of Autoquote, specialists were required to manually disseminate the inside quote (best bid and ask). The introduction of Autoquote replaced this process with a new automated quote whenever the NYSE limit order book changed. Hendershott, Jones and Menkveld (2011) documented that the introduction of Autoquote on NYSE narrows bid-ask spreads, particularly for large capitalisation stocks, attributed to an increase in algorithmic trading. Autoquote was phased in for different stocks over the period from 29 January 2003 to 27 May 2003. For our analysis, we use 1-week prior to the initial phasing in of Autoquote for the pre-event window and 1-week post the full introduction of Autoquote for the post-event window.

3.2.3. *Tick-Size Pilot.* The Tick-Size Pilot (TSP) was launched by the SEC on 3 October 2016 to assess the impact of an increase in tick size on market quality. In this study, we focus on stocks that underwent the tick size quote (Q) rule in which pilot stocks were quoted in \$0.05 increments but traded in \$0.01 increments.[5] Previous literature has documented the TSP Q rule to widen bid-ask spreads for pilot stocks (Chung, Lee & Rosch, 2020), due to the binding nature of the tick size. Eligible stocks for the Q rule were phased in between 3 October 2016 and 17 October 2016. Therefore, we use the 1-week prior to 3 October 2016 as the pre-event window and the 1-week after 17 October 2016 as the post-event window.

3.3. Descriptive stats. Table 1 reports summary statistics on the bid-ask spread and volume before and after the implementation of Decimalisation, Autoquote, and the TSP. Bid-ask spreads narrow in response to Decimalisation and Autoquote, with a simultaneous increase in volume. These results are in line with previous literature examining liquidity around Decimalisation (Bessembinder, 2003) and Autoquote (Hendershott, Jones & Menkveld, 2011). On the contrary, bid-ask spreads widen and volumes decline in response to the TSP Q rule, consistent with Chung, Lee and Rosch (2020).

[Table 1 here]

3.4. Simple gains from trade calculation.

3.4.1. *Actual volumes.* Using the spread and volume estimates of each stock (averages are documented in Table 2), we can estimate the trader gains (Eq. 6) and market maker revenues (7) for

[5] We focus our analysis on the TSP Q rule only because Chung, Lee and Rosch (2020) finds that ”both the T and TA rules do not affect quoted spreads.” (p. 889).

each stock. Upper and lower bounds for trader gains, summed across all stocks, are reported in Table 2. Upper bounds correspond to pure limit-order markets with no traders acting purely as market makers. In contrast, the lower bound represents a pure dealer market, where a dealer acts as the counterparty to every trade.

[Table 2 here]

Table 2 documents annual increases in trader gains from trade ranging from $143 million to $286 million for NYSE decimalisation, $471 million to $942 million for NASDAQ Decimalisation, and $317 million to $634 million for Nyse Autoquote. On the contrary, decreases in gains from trade ranged from -$2 million to -$4 million for NYSE TSP Q and -$1.9 million to -$3.8 million for Nasdaq TSP. Market maker revenue declined from Decimalisation but gained from Autoquote and the TSP Q rule.

The benefit of this approach is in its simplicity: gains from trade can be calculated for each stock with only four data points (pre and post volume and bid-ask spreads). However, it suffers from three significant shortcomings. First, Eq. (6) implicitly assumes a uniform private value distribution, which may not hold empirically. Second, post spreads and volumes are not observable *ex-ante*. Finally, volume changes can be contaminated by other simultaneous changes occurring as a stock transitions to a new market design that are not related to bid-ask spread. The subsequent analysis documents an approach to overcome these shortcomings.

3.4.2. *Predicted volumes.* Rather than actual changes to spread and volume, predicted volume uses historical data to estimate the relationship between volume and spreads. This somewhat alleviates the latter two concerns since volume is no longer required to be *ex-ante* observable and the relationship between bid-ask spreads and volume is estimated in the 6-months prior to the market design change, avoiding contamination. Predicted volume is obtained using Ordinary Least Squares (OLS) for each stock i using bid-ask spreads pre and post market design change as inputs. Gains from trade estimates using predicted volume are reported in Table 3.

[Table 3 here]

Overall conclusions are not significantly impacted, however, there are two minor effects on gains from trade estimates. First, Autoquote positive market maker revenue switching to negative can occur because changes in spread are less sensitive to changes in volume when using past data. This can occur through increased participation by algorithmic traders, which is not driven soley by tighter spreads. Second, market maker revenues are higher using predicted volume for TSP Q: volume is actually less sensitive using historical data, as market maker revenues increased more from wider spreads.

4. Estimating private values and gains from trade

4.1. Estimating trader private values.

4.1.1. *Empirical Strategy.* Trader private values are not observable but can be estimated using traders aggregate demand curve (e.g., Eq. 3). Using such a demand curve enables us to calculate, for example, the total trade volume with a private value within 5 and 10 basis points of the midquote. That is, if the bid-ask spread narrows from 20 basis points to 10 basis points, any additional volume reflects traders that can now transact and benefit from gains from trade.

Estimating the demand curve is not straightforward because the functional form of the relationship between bid-ask spreads and volume depends on the distribution of trader private values which is unobservable. Our empirical approach relies on discretising the sample into 10 bins sorted by relative bid-ask spread and estimating a linear model within each bin.[6] We can effectively divide the full sample into 10 sub-samples and estimate the following model specification using OLS for each stock i:

$$V_{it} = \beta_{i1} + \beta_{i2}S_{it} + \sum_{j=1}^{9} \gamma_{ij}d_{ij} + \sum_{j=1}^{9} \delta_{ij}d_{ij}S_{it} + \epsilon_{it} \tag{7}$$

The intercept in the first segment is β_{i1}, in the second segment $\beta_{i1} + \gamma_{i1}$, in the third segment $\beta_{i1} + \gamma_{i1} + \gamma_{i2}$ etc. The slope in the first segment is β_{i2}, in the second segment $\beta_{i2} + \delta_{i1}$, in the third segment $\beta_{i2} + \delta_{i1} + \delta_{i2}$ etc. Let the threshold points (bid-ask spread deciles), known henceforth as knots, be given by $\mu_i = (\mu^*_{i1}, \mu^*_{i2}, ..., \mu^*_{i9})$ where $d_{iz} = 1$ if $S_{it} \geq \mu^*_{iz}$.

A shortcoming of the model presented in Eq. (7) is that the function is not necessarily continuous at the threshold points (bid-ask spread deciles). A remedy is to use a spline regression that makes the function piecewise continuous, thereby joining at the knots (Greene, 2002). This requires restrictions on model coefficients so that

$$\left(\beta_{i1} + \sum_{j=1}^{z-1} \gamma_{ij}\right) + \left(\beta_{i2} + \sum_{j=1}^{z-1} \delta_{ij}\right) t^*_{iz} = \left(\beta_{i1} + \sum_{j=1}^{z} \gamma_{ij}\right) + \left(\beta_{i2} + \sum_{j=1}^{z} \delta_{ij}\right) t^*_{iz} \tag{8}$$

[6] Robustness tests in Section 6.2 vary the number of bins and this does not have a significant impact on conclusions.

for $j = 1, 2, ..., 9$. Collecting like terms and inserting these restrictions into Eq. (7) obtains

$$V_{it} = \beta_{i1} + \beta_{i2} S_{it} + \sum_{j=1}^{9} \delta_{ij} d_{ij} \left(S_{it} - \mu_{ij}^{*} \right) + \epsilon_{it} \tag{9}$$

Coefficients can be estimated with OLS specifying independent variables as S_{it}, $S_{it} - \mu_{i1}^{*}$ if $S_{it} \geq \mu_{i1}^{*}$ and zero otherwise, $S_{it} - \mu_{i2}^{*}$ if $S_{it} \geq \mu_{i2}^{*}$ and zero otherwise,..., $S_{it} - \mu_{i9}^{*}$ if $S_{it} \geq \mu_{i9}^{*}$ and zero otherwise.

Figure 2 illustrates the construction of the demand curve for NVIDIA on the NASDAQ.

[Figure 2 here]

Panel A plots the relationship between bid-ask spreads and dollar volume using the spline regression (Eq. 9) and Panel B plots the relationship between bid-ask spreads and dollar volume using OLS regressions. The general decreasing convex association between dollar volume and bid-ask spreads is common, regardless of the approach. The main difference is that the standard OLS regressions exhibit discontinuities at each bid-ask spread decile.

4.1.2. *Results.* Table 4 reports the dollar volume associated with private values between each decile of the bid-ask spread range across the universe of NYSE and Nasdaq stocks respectively. These results are calculated using the spline regression (Eq. 9) for each stock separately and then aggregated for the market. Dollar volumes at tighter spreads are obtained from extrapolating the spline regression to zero bid-ask spread using the gradient of the demand curve in the lowest bid-ask spread segment. Dollar volume at wider spreads is the predicted volume using Eq. (9) at the widest bid-ask spread during the sample period.

[Table 4 here]

Panel A of Table 4 documents that more dollar volume is extracted from a reduction in bid-ask spreads (per bps) as bid-ask spreads get tighter per annum from January 1996 to December 2022. For example, reducing bid-ask spreads from 0.06 bps to zero increases per annum volume by $302 million per stock on NYSE. This contrasts with a reduction in bid-ask spread from 89 bps to 56 bps increasing per annum volume by only $170 million per stock. On the face of it, this suggests more traders have private values concentrated around the midquote. However, the full sample results are, to a degree, influenced by increases in average trading volumes and decreases in average bid-ask spreads through time (Roll & Subrahmanyam, 2010).

Panel B of Table 4 documents that for the most recent year in our analysis (2022), these main insights hold. Specifically, dollar volume per bps reduction in bid-ask spread is higher for tighter bid-ask spreads. In this case, reducing bid-ask spreads from 0.06 bps to zero would result in an increase in volume of $1.653 billion per annum on NYSE.[7]

4.2. Gains from trade due to market design change. We calculate gains from trade for each stock by estimating the demand curve using six months of data prior to the implementation of Decimalisation, Autoquote, and TSP Q rules.

[Table 5 here]

Our main conclusions remain consistent with the simpler OLS approach presented in Table 3: the signs of trader gains and market maker revenue changes are unchanged. However, allowing for non-linearity in the demand curve slightly alters the magnitude of gains from trade. As shown in Table 5, estimated trader gains range from $39 million to $78 million for NYSE Decimalisation and from $1.16 billion to $2.31 billion for NASDAQ Decimalisation. For NYSE Autoquote, gains range from $90 million to $180 million. TSP Q gains range from -$53 million to -$106 million on NASDAQ and NYSE. These ranges reflect two extreme assumptions: a pure dealer market where each trade involves a market maker (first number) versus a pure limit order market with no market maker counterparties (second number).

Trader gains consist of two components: (1) the additional surplus to existing traders from a narrower bid-ask spread, and (2) the surplus gained by traders who begin participating once spreads narrow enough to make trading worthwhile. Under these assumptions, traders benefit from NYSE Decimalisation, NASDAQ Decimalisation, and NYSE Autoquote. In contrast, trader gains are negative for both NYSE and NASDAQ TSP Q.

The impact to market maker revenues depend on how spreads and trading volume respond to the market design change. Market makers revenue increases from a *reduction* in the bid-ask spread only if the product of spread and volume increases, i.e., if $S_{post}V_{post} > S_{pre}V_{pre}$. Otherwise, they revenue increases from a *widening* of spreads. Maximum market maker revenue losses - corresponding to a pure dealer market - are $5 million for NYSE Decimalisation, $1.19 billion for NASDAQ Decimalisation, and $72 million for NYSE Autoquote. Maximum revenue gains (pure dealer market) are $66 million for NYSE TSP Q and $78 million for NASDAQ TSP Q. In general,

[7]Lower volume as bid-ask spreads decline for wider bid-ask spread percentiles occurs due to positive spread-volume relationships. This occurs for illiquid stocks that are less actively monitored or traded.

market makers incur revenue losses following reforms that narrowed spreads (Decimalisation, Autoquote) and gain following reforms that widened them (TSP Q). This pattern suggests that bid-ask spreads and volume on NYSE and NASDAQ are not sufficiently responsive to enable market maker revenues to increase from spread reductions.

5. Extensions

5.1. Total gains from trade. Total gains from trade are estimated by extrapolating the demand curve to the point of zero dollar volume, using the slope from the segment with the widest bid-ask spreads (between the 90th and 100th percentiles). Zero dollar volume corresponds to a fully prohibitive spread, where no traders have private values further from the midquote to justify trading. For each stock, we calculate realized gains from trade using the average bid-ask spread over the sample period as a representative measure.[8] We also compute maximum gains from trade, defined as the gains that would be realized in the hypothetical case of zero bid-ask spreads. The difference between actual and maximum gains reflects the deadweight loss due to the bid-ask spread. Table 6 presents aggregate gains from trade across the NYSE and NASDAQ between 1996 and 2022.

Total gains from trade are higher when a) average bid-ask spreads are narrower, allowing more trades to occur, and b) traders' private values lie further from the prevailing bid or ask, increasing the surplus from each trade. Maximum gains from trade increase not only when private values are more dispersed, but also when trading volume is more sensitive to reductions in the spread at tight spread ranges. Conversely, the proportion of gains realized (i.e., actual gains as a share of maximum gains) is highest when spreads are tight and volume is relatively insensitive to marginal spread changes at higher spread ranges, indicating fewer foregone trades.

[Table 6 here]

Table 6 shows that, on average, 89% of maximum gains from trade were realized on NASDAQ and 91% on the NYSE over the 1996–2022 period.[9] Average annual gains from trade totaled $172.44 billion on NASDAQ and $142.63 billion on NYSE. The proportion of maximum gains achieved peaked in 2001 and hit its lowest point in 2010 on both exchanges. This suggests that in 2001, either spreads were more conducive to trading or private values were more dispersed, whereas in 2010,

[8] Total gains from trade comprise both trader surplus and market maker revenues. Even if market maker revenues reflect only compensation for providing liquidity rather than net economic gains, they are essential for market functioning and should not be considered deadweight loss.

[9] The proportion of maximum gains on NYSE and NASDAQ are comparable to estimates by Hollifield et al. (2006) for three stocks on the Vancouver Stock Exchange.

trading activity was more constrained—possibly due to greater clustering of private values around the midquote or heightened sensitivity of volume to spreads at tighter spreads. Notably, 2010 also marked the lowest total gains from trade on both exchanges, providing support that trader private values were less dispersed. In contrast, the highest gains were observed in 2018 for NASDAQ and in 2008 for NYSE.

5.2. Characteristics. This section examines which stock-level characteristics influence the sensitivity of trading volume to changes in bid-ask spreads using a two-stage regression approach. A higher elasticity of volume with respect to the bid-ask spread reflects greater potential gains from trade when spreads narrow-or greater losses when they widen. This has important implications for policymakers interested in targeting liquidity improvements across different types of stocks.

In the first stage, we estimate daily spread-volume sensitivity by regressing 5-minute interval dollar trading volume on the bid-ask spread for each stock i and each day t using OLS:

$$V_{itz} = \gamma_{it} + \beta_{it} S_{itz} + \epsilon_{itz} \tag{10}$$

where V_{itz} denotes dollar volume, S_{itz} is the bid-ask spread, and z indexes 5-minute intervals. The estimated coefficients $\hat{\beta}_{it}$ capture the responsiveness of volume to the spread on a given day and stock.

In the second stage, we regress the estimated spread-volume sensitivities $\hat{\beta}_{it}$ on a set of observable stock characteristics:

$$\begin{aligned}\hat{\beta}_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 MarketCap_{it} + \beta_3 Volatility_{it} + \beta_4 AlgoTrading_{it} \\ + \beta_5 TradeSize_{it} + \beta_6 Nasdaq_i + \mu_i + \tau_{t \in T} + \epsilon_{it}\end{aligned} \tag{11}$$

Here, S_{it} is the average daily bid-ask spread; *MarketCap* is computed as the number of shares outstanding times the share price (in billions); *Volatility* is measured as the absolute daily return; *AlgoTrading* is proxied by the ratio of quote messages to trades; *TradeSize* is the average number of shares per trade; and *Nasdaq* is a dummy variable equal to 1 if a stock is listed on NASDAQ. We include stock fixed effects (μ_i) and month fixed effects ($\tau_{t \in T}$) to control for unobserved heterogeneity across stocks and over time where T indexes each month.

Table 7 reports estimated coefficients and t-statistics across four specifications: (1) no fixed effects; (2) stock fixed effects only; (3) month fixed effects only; and (4) both stock and month fixed effects. Columns (5)–(8) repeat the analysis using spread-volume sensitivities estimated at a monthly frequency.[10]

[Table 7 here]

Table 7 yields three key findings. First, stocks with lower market capitalisation exhibit more negative spread-volume sensitivity, even after controlling for time and stock-specific effects. This suggests that changes in the bid-ask spread have a more pronounced effect on gains from trade (in basis points) for smaller stocks. Intuitively, small-cap stocks resemble luxury goods-displaying greater elasticity-while large-cap stocks behave more like necessities, due to their presence in leading indices.

Second, higher volatility is associated with more negative spread-volume sensitivity. Thus, in more volatile markets or for more volatile stocks, changes in the spread have a larger impact on trading volume. While it is well-established that volatility correlates positively with both volume (Karpoff, 1987) and bid-ask spreads (Stoll, 1978), its role in amplifying spread-volume sensitivity is less explored. A possible explanation is that heightened price risk in volatile markets amplify trader sensitivity to the implicit transaction cost of the bid-ask spread.

Third, stocks with higher levels of algorithmic trading, particularly evident in the daily analysis, display more negative spread-volume sensitivity. This implies that algorithmic trading may enhance the responsiveness of volume to changes in the bid-ask spread. A plausible mechanism is that algorithmic traders, with their superior monitoring and reaction speed, are better positioned to adjust trading intensity as spreads change.

6. Robustness

6.1. Control variables and endogeneity. Our main gains from trade analysis in Section 4.2 is potentially affected by two key concerns: (a) omitted variable bias, and (b) endogeneity between spreads and volumes. To partially address these issues, we implement a two-stage approach to obtain predicted volumes as a robustness check. In the first stage, we estimate predicted bid-ask spreads by regressing current spreads on their lagged values:

$$S_{it} = \beta_{i1} + \beta_{i2} S_{i,t-1} + \epsilon_{it} \tag{12}$$

[10] Spread-volume sensitivities ($\hat{\beta}_{it}$) are scaled by 100 million for readability.

This stage helps mitigate endogeneity by using lagged bid-ask spreads as an instrument for current bid-ask spreads. In the second stage, we use the fitted values $\hat{S}_{it}$ from Eq. 12 and include additional control variables to reduce omitted variable bias, modifying Eq. (9) as follows:

$$V_{it} = \beta_{i1} + \beta_{i2}\hat{S}_{it} + \sum_{j=1}^{9} \delta_{ij} d_{ij} \left(\hat{S}_{it} - \mu^*_{ij}\right) + \beta_{i3} Volatility_{it} + \beta_{i4} MarketCap_{it} \tag{13}$$
$$+ \beta_{i5} Return_{it} + \sum_{j=6}^{7} \beta_{ij} TimeofDay_{itj} + \sum_{j=8}^{11} \beta_{ij} DayofWeek_{itj} + \tau_{t\in T} + \epsilon_{it}$$

Here, *Return* denotes 5-minute simple returns,[11] *DayofWeek* refers to dummy variables capturing day-of-week effects on volume, and *TimeofDay* indicates dummy variables for the first and last hour of trading, when volume is typically elevated (Chung, Van Ness & Van Ness, 1999). *Volatility* and *MarketCap* are defined in Eq. (11). The positive relationships between both market capitalisation and volume, and volatility and volume,,are well-documented (e.g., Choi, 2019; Karpoff, 1987). We present gains-from-trade estimates using predicted volumes from Eq. (13) in Table 8.

[Table 8 here]

Our main findings from Table 5 remain broadly unchanged. However, some nuanced differences emerge. For both exchanges, market maker revenues increase following Decimalisation. On NASDAQ, this comes at the expense of trader gains, suggesting that predicted volume before Decimalisation is lower due to weaker spread-volume sensitivity (an inelastic demand curve) at spreads below S_{pre}. Conversely, on the NYSE, trader gains also rise, indicating a greater difference between predicted pre- and post-Decimalisation volumes, consistent with a more elastic demand curve between S_{pre} and S_{post}. For NYSE Autoquote, the pattern mirrors NASDAQ Decimalisation: gains shift from traders to market makers. A similar affect is observed for both exchanges in response to TSP *Q*: market maker revenues fall while trader gains increase in response to wider average bid-ask spreads.

6.2. Number of knots in demand curve estimation. Our main analysis calculates gains from trade using a spline regression with 10 knots, chosen somewhat arbitrarily. A high amount of knots increases the precision of the demand curve at the potential cost of overfitting. Table 9 documents gains from trade due to Decimalisation, Autoquote, and TSP *Q* using 5, 10, and 15 knots.

[11]Some studies find a stronger volume-volatility relationship for positive returns (e.g., Brailsford, 1996). Therefore, we include non-absolute returns as a complementary control to volatility.

[Table 9 here]

The number of knots does not have a significant impact on conclusions made in the paper.

7. Conclusion

This paper develops a novel and practical approach to estimating gains from trade in financial markets. Our methodology is simple, transparent, and feasible for large-scale implementation, overcoming the limitations of earlier studies such as Hollifield et al. (2006) that rely heavily on structural assumptions and are often computationally intensive. Our approach leverages the inverse relationship between trading volumes and bid-ask spreads to construct a demand curve reflecting the distribution of trader private values for a stock, providing a direct and flexible measure of gains from trade.

We demonstrate the utility of our approach through three prominent market design changes: Decimalisation, Autoquote, and the Tick Size Pilot (TSP Q rule). In each case, we show how our method can be used *ex-post* to evaluate the impact of reforms on market participants' welfare. Importantly, our method also offers the potential to inform *ex-ante* policy assessments, provided regulators have a forecast of the policy's impact on bid-ask spreads.

A distinguishing feature of our approach is the range-based nature of the gains-from-trade estimates, which reflect the uncertainty surrounding the proportion of traders acting as market makers in limit order markets. Developing robust estimates of this proportion is an important direction for future research that would further refine our welfare estimates and enhance their value for policy evaluation.

Looking ahead, an important extension is to explore the economic forces that shape both total and maximum gains from trade in a market. Because these gains tend to be larger when trader private values are more dispersed relative to top of book bid and ask quotes, future work could examine whether observable proxies for disagreement - such as volatility or dispersion in analyst forecasts - can explain cross-market or temporal variation in gains from trade.

References

Bessembinder, H. (2003), 'Trade execution costs and market quality after decimalization', *Journal of Financial and Quantitative Analysis* **38**(4), 747–777.

Brailsford, T. J. (1996), 'The empirical relationship between trading volume, returns and volatility', *Accounting & Finance* **36**(1), 89–111.

Brogaard, J., Hagströmer, B., Nordén, L. and Riordan, R. (2015), 'Trading fast and slow: Colocation and liquidity', *The Review of Financial Studies* **28**(12), 3407–3443.

Cason, T. N. and Friedman, D. (1996), 'Price formation in double auction markets', *Journal of Economic Dynamics and Control* **20**(8), 1307–1337.

Choi, H. M. (2019), 'Market uncertainty and trading volume around earnings announcements', *Finance Research Letters* **30**, 14–22.

Chung, K. H., Lee, A. J. and Rösch, D. (2020), 'Tick size, liquidity for small and large orders, and price informativeness: Evidence from the tick size pilot program', *Journal of Financial Economics* **136**(3), 879–899.

Chung, K. H., Van Ness, B. F. and Van Ness, R. A. (1999), 'Limit orders and the bid–ask spread', *Journal of Financial Economics* **53**(2), 255–287.

Chung, K. H., Van Ness, B. F. and Van Ness, R. A. (2004), 'Trading costs and quote clustering on the nyse and nasdaq after decimalization', *Journal of Financial Research* **27**(3), 309–328.

DeFusco, A. A., Tang, H. and Yannelis, C. (2022), 'Measuring the welfare cost of asymmetric information in consumer credit markets', *Journal of Financial Economics* **146**(3), 821–840.

Einav, L., Finkelstein, A. and Cullen, M. R. (2010), 'Estimating welfare in insurance markets using variation in prices', *The Quarterly Journal of Economics* **125**(3), 877–921.

Foley, S., Kwan, A., Philip, R. and Ødegaard, B. A. (2022), 'Contagious margin calls: How covid-19 threatened global stock market liquidity', *Journal of Financial Markets* **59**, 100689.

Glosten, L. R. and 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.

Glosten, L. R. and Putnins, T. J. (2016), 'Welfare costs of informed trade', *Working Paper, Columbia Business School and UTS Business School* .

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

Goldstein, M. A. and Hotchkiss, E. S. (2020), 'Providing liquidity in an illiquid market: Dealer behavior in us corporate bonds', *Journal of Financial Economics* **135**(1), 16–40.

Greene, W. H. (2002), *Econometric Analysis*, Prentice Hall.

Harberger, A. C. (1966), *Efficiency effects of taxes on income from capital*.

Hausman, J. A. (1981), 'Exact consumer's surplus and deadweight loss', *The American Economic Review* **71**(4), 662–676.

Hendershott, T., Jones, C. M. and Menkveld, A. J. (2011), 'Does algorithmic trading improve liquidity?', *The Journal of Finance* **66**(1), 1–33.

Hollifield, B., Miller, R. A., Sandås, P. and Slive, J. (2006), 'Estimating the gains from trade in limit-order markets', *The Journal of Finance* **61**(6), 2753–2804.

Kagel, J. H. (2004), Double auction markets with stochastic supply and demand schedules: Call markets and continuous auction trading mechanisms, *in* 'Advances in understanding strategic Behaviour: Game Theory, Experiments and Bounded Rationality', Springer, pp. 181–208.

Karpoff, J. M. (1987), 'The relation between price changes and trading volume: A survey', *Journal of Financial and Quantitative Analysis* **22**(1), 109–126.

Kyle, A. S. (1985), 'Continuous auctions and insider trading', *Econometrica* **15**(35), 1315–1335.

Liu, H. and Wang, Y. (2016), 'Market making with asymmetric information and inventory risk', *Journal of Economic Theory* **163**, 73–109.

Riordan, R. and Storkenmaier, A. (2012), 'Latency, liquidity and price discovery', *Journal of Financial Markets* **15**(4), 416–437.

Roll, R. and Subrahmanyam, A. (2010), 'Liquidity skewness', *Journal of Banking & Finance* **34**(10), 2562–2571.

Stoll, H. R. (1978), 'The supply of dealer services in securities markets', *The Journal of Finance* **33**(4), 1133–1151.

Tables

Table 1: Summary statistics
This table reports average % bid-ask spreads and volume (millions of dollars) for effected stocks during intervals on either side of decimalisation on NYSE (1778 stocks) and NASDAQ (2260 stocks), Autoquote on NYSE (2408 stocks), and Tick Size Pilot Q rule on NYSE (125 stocks) and NASDAQ (254 stocks). For each market design change, the pre event period ends at the beginning of the onboarding process and the post event period starts at the end of the onboarding process. Decimalisation and Autoquote impact the universe of stocks on NASDAQ and NYSE (only Autoquote), while the Tick Size Pilot only includes a subset of stocks impacted by the quoting (Q) rule change. The NYSE decimalisation pre-event period is prior to 28 August 2000 and the post-event period is after 29 January 2001. The NASDAQ decimalisation pre-event period is prior to 12 March 2001 and the post-event period is after 9 April 2001. For Autoquote, the pre-event period is prior to 29 January 2003 and the post-event period is after 27 May 2003. For the TSP Q rule, the pre event window is 1 week prior to 3 October 2016 and the post-event window is after 17 October 2016.

Table 1. Summary Statistics

	Pre Spread (%)	Post Spread (%)	Pre Volume ($ Millions)	Post Volume ($ Millions)
NYSE - decimalisation	2.86	2.67	11.00	21.09
NASDAQ - decimalisation	1.67	1.54	12.69	17.43
NYSE - Autoquote	1.13	0.92	14.70	16.87
NYSE - TSP Q	0.48	0.71	6.04	5.51
NASDAQ - TSP Q	0.99	1.20	3.12	2.91

Table 2: Simple gains from trade (actual volumes)
This table reports annual gains from trade estimates in billions ($) resulting from Decimalisation, Autoquote, and the TSP Rule Q. Gains from trade are reported separately for NASDAQ stocks and NYSE stocks (only NYSE for Autoquote). Average bid-ask spreads and dollar volumes are obtained for each stock in the 1-week before (pre) and 1-week after (post) each market design change. Incremental trader gains are calculated as:

$$\Delta TraderGains = (1 + MM)\left(\frac{1}{2}\left(S_{pre} - S_{post}\right)V_{pre} + \frac{1}{4}\left|\left(S_{pre} - S_{post}\right)\right|\left(V_{post} - V_{pre}\right)\right).$$

where $MM = 0$ (pure dealer market) corresponds to the lower bound and $MM = 1$ (no intermediation) corresponds to the upper bound. Incremental market maker revenue gains are calculated as:

$$\Delta MMRevenue = \frac{1}{2}(1 - MM) \times \left(S_{post}\left(V_{post} - V_{pre}\right) - \left(S_{pre} - S_{post}\right)V_{pre}\right).$$

These estimates are summed for each stock.

Table 2. Simple gains from trade (actual volumes)

	Traders' gains from trade ($billions)	Market makers' revenue gain ($billions)
NYSE - decimalisation (lower)	0.143	-0.031
NYSE - decimalisation (upper)	0.286	-
NASDAQ - decimalisation (lower)	0.471	-0.160
NASDAQ - decimalisation (upper)	0.942	-
NYSE - Autoquote (lower)	0.317	0.175
NYSE - Autoquote (upper)	0.634	-
NYSE - TSP Q (lower)	-0.020	0.0056
NYSE - TSP Q (upper)	-0.040	-
NASDAQ - TSP Q (lower)	-0.019	0.0085
NASDAQ - TSP Q (upper)	-0.038	-

Table 3: Simple gains from trade (predicted volumes)
This table reports annual gains from trade estimates in billions ($) resulting from Decimalisation, Autoquote, and the TSP Rule Q. Gains from trade are reported separately for NASDAQ stocks and NYSE stocks (only NYSE for Autoquote). Average bid-ask spreads are obtained for each stock in the 1-week before (pre) and 1-week after (post) each market design change. Pre- and Post-predicted volume is obtained by regressing volume on bid-ask spreads in the 6-months before each market design change and using pre and post bid-ask spreads as input. Incremental trader gains are calculated as:

$$\Delta TraderGains = (1 + MM)\left(\frac{1}{2}\left(S_{pre} - S_{post}\right)\hat{V}_{pre} + \frac{1}{4}\left|\left(S_{pre} - S_{post}\right)\right|\left(\hat{V}_{post} - \hat{V}_{pre}\right)\right).$$

where $MM = 0$ (pure dealer market) corresponds to the lower bound and $MM = 1$ (no intermediation) corresponds to the upper bound. Incremental market maker revenue gains are calculated as:

$$\Delta MMRevenue = \frac{1}{2}(1 - MM) \times \left(S_{post}\left(\hat{V}_{post} - \hat{V}_{pre}\right) - \left(S_{pre} - S_{post}\right)\hat{V}_{pre}\right).$$

These estimates are summed for each stock.

Table 3. Simple gains from trade (predicted volumes)

	Traders' gains from trade ($billions)	Market makers' revenue gain ($billions)
NYSE - decimalisation (lower)	0.06	-0.08
NYSE - decimalisation (upper)	0.12	-
NASDAQ - decimalisation (lower)	1.29	-1.16
NASDAQ - decimalisation (upper)	2.58	-
NYSE - Autoquote (lower)	0.09	-0.094
NYSE - Autoquote (upper)	0.18	-
NYSE - TSP Q (lower)	-0.036	0.042
NYSE - TSP Q (upper)	-0.072	-
NASDAQ - TSP Q (lower)	-0.047	0.043
NASDAQ - TSP Q (upper)	-0.094	-

Table 4: Private value distributions
This table reports the increase in dollar volume associated with a reduction in bid-ask spread from the 100%ile to 90%ile, 90%ile to 80%ile etc. for the entire distribution of bid-ask spreads on the NYSE and NASDAQ. Results are reported in millions ($) and scaled by the number of stocks. The number of stocks on the Nasdaq are 10487 in the full sample and 2767 in 2022. The number of stocks on the NYSE are 7196 in the full sample and 2141 in 2022. Tight refers to the dollar volume associated with reducing the bid-ask spread from the minimum value during the sample period to zero and Wide refers to the dollar volume associated with the maximum bid-ask spread in the sample period. Spread range is the bid-ask spread range for a given distribution that corresponds to the dollar volume. The dollar volume increase for a given reduction in bid-ask spread is calculated using a spline regression for each stock in the sample with 10 knots at each decile of bid-ask spread.

Table 4. Private value distributions

	Tight	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Wide
Panel A. Full Sample												
NYSE												
Spread Upper Bound (bps)	0.06	4	8	14	23	36	56	89	148	303	20000	$+\infty$
Dollar Volume	302	9354	1654	1024	669	394	242	170	98	74	40	3542
Dollar Volume (per bps)	5041	3141	621	492	289	169	68	25	7	4	0.04	
NASDAQ												
Spread Upper Bound (bps)	0.23	9	20	34	54	82	126	195	317	606	20000	$+\infty$
Dollar Volume	446	5668	1268	757	542	374	268	186	163	114	52	1598
Dollar Volume (per bps)	1925	655	119	53	27	13	6	3	1	0.4	0.003	
Panel B. 2022												
NYSE												
Spread Upper Bound (bps)	0.2	8	16	26	39	58	85	131	213	414	6667	$+\infty$
Dollar Volume	4306	43220	3620	381	-294	-631	-753	-767	-531	-97	33	25360
Dollar Volume (per bps)	27880	14029	1789	153	-97	-149	-123	-76	-26	-2	0.01	
NASDAQ												
Spread Upper Bound (bps)	0.26	6	11	18	27	39	59	91	157	305	11773	$+\infty$
Dollar Volume	7158	38106	715	-677	-824	-406	-151	-40	37	27	17	23158
Dollar Volume (per bps)	27259	6992	147	-99	-95	-32	-8	-1	1	0.2	0.001	

Table 5: Gains from trade with estimated demand curve

This table reports gains from trade estimates in billions ($) resulting from Decimalisation, Autoquote, and the TSP Rule Q on NASDAQ and NYSE. Average bid-ask spreads are obtained for each stock in the 1-week before (pre) and 1-week after (post) each market design change. A linear spline regression with 10 knots at bid-ask spread deciles is estimated for each stock to obtain the demand curve. Trader gains from trade due to a market design change is the area under the curve from pre to post spreads, which is cumulated along each linear spline as:

$$\Delta TraderGains = (1 + MM)\left(\frac{1}{2}\left(S_{pre} - S_{post}\right)\hat{V}_{pre} + \frac{1}{4}\left|\left(S_{pre} - S_{post}\right)\right|\left(\hat{V}_{post} - \hat{V}_{pre}\right)\right).$$

where $MM = 0$ (pure dealer market) corresponds to the lower bound and $MM = 1$ (no intermediation) corresponds to the upper bound. Incremental market maker revenue gains are calculated as:

$$\Delta MMRevenue = \frac{1}{2}(1 - MM) \times \left(S_{post}\left(\hat{V}_{post} - \hat{V}_{pre}\right) - \left(S_{pre} - S_{post}\right)\hat{V}_{pre}\right).$$

These estimates are summed for each stock.

Table 5. Gains from trade with estimated demand curve

	Traders' gains from trade ($billions)	Market makers' revenue gain ($billions)
NYSE - Decimalisation (lower)	0.039	-0.005
NYSE - Decimalisation (upper)	0.078	-
NASDAQ - Decimalisation (lower)	1.157	-1.191
NASDAQ - Decimalisation (upper)	2.314	-
NYSE - Autoquote (lower)	0.090	-0.072
NYSE - Autoquote (upper)	0.180	-
NYSE - TSP Q (lower)	-0.053	0.066
NYSE - TSP Q (upper)	-0.106	-
NASDAQ - TSP Q (lower)	-0.053	0.078
NASDAQ - TSP Q (upper)	-0.106	-

Table 6: Total and maximum gains from trade
This table reports annual gains from trade estimates in billions ($) from 1996 to 2022. Gains from trade are reported separately for NASDAQ stocks and NYSE stocks. Total gains from trade is the aggregate gains from trade for each stock assuming the average bid-ask spread in the relevant period. Maximum gains from trade is the aggregate gains from trade for each stock assuming a zero-bid ask spread.

Table 6. Total and maximum gains from trade

Year	NASDAQ			NYSE		
	Total Gains ($billions)	Maximum Gains ($billions)	Gains Proportion	Total Gains ($billions)	Maximum Gains ($ billions)	Gains Proportion
1996	170.96	205.13	0.83	72.15	96.31	0.75
1997	219.99	233.37	0.94	82.92	87.78	0.94
1998	224.60	244.90	0.92	111.72	118.70	0.94
1999	321.16	355.21	0.90	155.89	166.13	0.94
2000	444.28	481.68	0.92	146.29	158.47	0.92
2001	389.66	396.03	0.98	125.12	125.34	0.998
2002	126.91	131.17	0.97	162.03	163.30	0.99
2003	47.78	53.70	0.89	55.27	58.12	0.95
2004	484.49	491.92	0.98	61.73	66.23	0.93
2005	41.67	48.00	0.87	46.04	49.74	0.93
2006	32.99	38.45	0.86	56.24	62.66	0.90
2007	55.60	65.68	0.85	107.42	118.92	0.90
2008	103.11	109.23	0.94	1048.05	1056.71	0.99
2009	39.27	45.42	0.86	113.93	121.88	0.93
2010	31.90	44.23	0.72	36.00	48.08	0.75
2011	34.53	46.47	0.74	80.84	94.35	0.86
2012	67.27	76.81	0.88	114.45	128.78	0.89
2013	67.26	79.18	0.85	55.31	62.97	0.88
2014	83.14	94.10	0.88	105.59	114.69	0.92
2015	76.06	92.93	0.82	69.09	81.56	0.85
2016	72.96	81.54	0.89	72.62	89.70	0.81
2017	108.06	123.11	0.88	71.41	84.90	0.84
2018	495.82	519.33	0.95	134.43	145.90	0.92
2019	147.35	159.94	0.92	132.27	143.64	0.92
2020	330.21	344.38	0.96	312.90	327.87	0.95
2021	204.82	232.91	0.88	140.99	156.58	0.90
2022	234.04	250.33	0.93	180.25	192.57	0.94
Average	172.44	186.86	0.89	142.63	152.66	0.91

Table 7: Characteristics analysis
This table reports coefficients and *t* statistics from an OLS regression of spread-volume sensitivity on bid-ask spreads (bps), market capitalisation (billions), volatility (absolute close to close returns), algorithmic trading (proxied as daily order to trade ratio), trade size and a dummy variable equal to 1 if a stock is listed on the NASDAQ. Spread-volume sensitivity is calculated for each stock on each day (columns 1 to 4) and each month (columns 5 to 8) by regressing dollar volume on bid-ask spreads using OLS and scaled by 100 million for readability. Independent variables are average daily values for the monthly spread-volume sensitivity specification. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level respectively. Standard errors are heteroskedasticity and autocorrelation consistent (HAC).

Table 7. Characteristics analysis

	Daily Betas				Monthly Betas			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spread	-239.65	-57.15	-90.09	282.46	-57.55	-115.09	-43.81	-135.85
	[-0.53]	[-0.34]	[-0.54]	[1.13]	[-0.26]	[-0.75]	[-0.53]	[-0.89]
Market Capitalisation	1.52***	2.18***	1.52***	2.10***	0.73***	1.17***	0.73***	1.15***
	[12.76]	[6.99]	[4.94]	[6.75]	[9.66]	[12.35]	[8.66]	[12.32]
Volatility	-386.38***	-417.26***	-367.31***	-383.24***	-247.32*	-287.58*	-250.63*	-299.67**
	[-2.72]	[-3.84]	[-3.59]	[-3.70]	[-1.92]	[-1.73]	[-1.67]	[-1.96]
Algo Trading	-0.02	-0.01**	-0.03***	-0.02**	-0.005	-0.005	-0.007	-0.006
	[-1.17]	[-2.11]	[-3.20]	[-2.45]	[-0.46]	[-1.37]	[-1.63]	[-1.38]
Trade Size	1.23	0.72	2.41	2.5	1.23	1.91	2.14	2.77
	[0.54]	[0.63]	[1.04]	[1.03]	[0.57]	[0.92]	[0.92]	[0.89]
Nasdaq	17.41**		16.17*		10.91**		10.84*	
	[2.56]		[1.93]		[2.48]		[1.93]	
Stock Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Month Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes

Table 8: Gains from trade after controlling for endogeneity and omitted variable bias
This table reports gains from trade estimates in billions ($) resulting from Decimalisation, Autoquote, and the TSP Rule Q on NASDAQ and NYSE. Predicted bid-ask spreads are estimated by regressing lagged bid-ask spreads on current bid-ask spreads. Average predicted bid-ask spreads are obtained for each stock in the 1-week before (pre) and 1-week after (post) each market design change. A linear spline regression with 10 knots at bid-ask spread deciles is estimated for each stock to obtain the demand curve while controlling for volatility (absolute daily return), market capitalisation, simple 5-minute returns, time of day dummies (first and last trading hours), day of week dummies and month fixed effects. Trader gains from trade due to a market design change is the area under the curve from pre to post spreads, which is cumulated along each linear spline as: Incremental trader gains are calculated as:

$$\Delta TraderGains = (1 + MM)\left(\frac{1}{2}\left(S_{pre} - S_{post}\right)\hat{V}_{pre} + \frac{1}{4}\left|\left(S_{pre} - S_{post}\right)\right|\left(\hat{V}_{post} - \hat{V}_{pre}\right)\right).$$

where $MM = 0$ (pure dealer market) corresponds to the lower bound and $MM = 1$ (no intermediation) corresponds to the upper bound. Incremental market maker revenue gains are calculated as:

$$\Delta MMRevenue = \frac{1}{2}(1 - MM) \times \left(S_{post}\left(\hat{V}_{post} - \hat{V}_{pre}\right) - \left(S_{pre} - S_{post}\right)\hat{V}_{pre}\right).$$

These estimates are summed for each stock.

Table 8. Gains from trade after controlling for endogeneity and omitted variable bias

	Traders' gains from trade ($billions)	Market makers' revenue gain ($billions)
NYSE - decimalisation (lower)	0.103	0.045
NYSE - decimalisation (upper)	0.206	-
NASDAQ - decimalisation (lower)	0.700	-0.479
NASDAQ - decimalisation (upper)	1.400	-
NYSE - Autoquote (lower)	0.032	-0.002
NYSE - Autoquote (upper)	0.064	-
NYSE - TSP Q (lower)	-0.013	0.018
NYSE - TSP Q (upper)	-0.026	-
NASDAQ - TSP Q (lower)	-0.036	0.022
NASDAQ - TSP Q (upper)	-0.072	-

Table 9: Gains from trade while varying number of knots
This table reports gains from trade estimates in billions ($) resulting from Decimalisation, Autoquote, and the TSP Rule Q. Average bid-ask spreads are obtained for each stock in the 1-week before (pre) and 1-week after (post) each market design change. A linear spline regression with 5 knots (Panel A), 10 knots (Panel B), and 15 knots (Panel C) at bid-ask spread deciles is estimated for each stock to obtain the demand curve. Trader gains from trade due to a market design change is the area under the curve from pre to post spreads, which is cumulated along each linear spline as:

$$\Delta TraderGains = (1 + MM)\left(\frac{1}{2}\left(S_{pre} - S_{post}\right)\hat{V}_{pre} + \frac{1}{4}\left|\left(S_{pre} - S_{post}\right)\right|\left(\hat{V}_{post} - \hat{V}_{pre}\right)\right).$$

where $MM = 0$ (pure dealer market) corresponds to the lower bound and $MM = 1$ (no intermediation) corresponds to the upper bound. Incremental market maker revenue gains are calculated as:

$$\Delta MMRevenue = \frac{1}{2}(1 - MM) \times \left(S_{post}\left(\hat{V}_{post} - \hat{V}_{pre}\right) - \left(S_{pre} - S_{post}\right)\hat{V}_{pre}\right).$$

These estimates are summed for each stock.

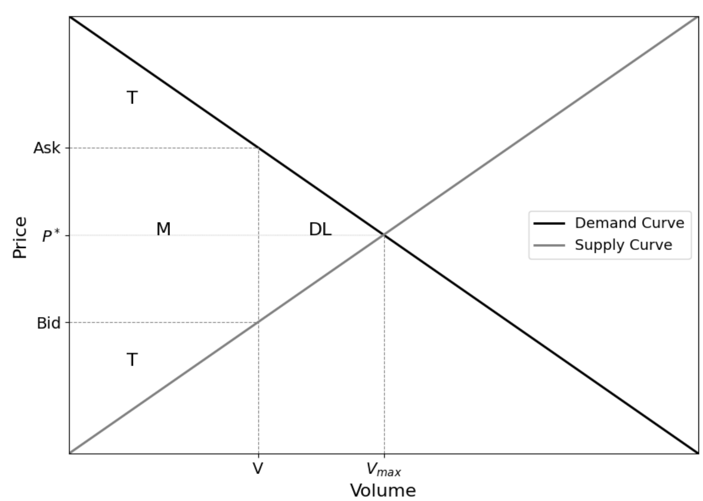
Table 9. Gains from trade while varying number of knots

	Traders' gains from trade ($billions)	Market makers' revenue gain ($billions)
Panel A. 5 Knots		
NYSE - Decimalisation (lower)	0.040	-0.050
NYSE - Decimalisation (upper)	0.080	-
NASDAQ - Decimalisation (lower)	1.102	-0.644
NASDAQ - Decimalisation (upper)	2.204	-
NYSE - Autoquote (lower)	0.090	-0.080
NYSE - Autoquote (upper)	0.180	-
NYSE - TSP Q (lower)	-0.045	0.052
NYSE - TSP Q (upper)	-0.090	-
NASDAQ - TSP Q (lower)	-0.043	0.048
NASDAQ - TSP Q (upper)	-0.086	-
Panel B. 10 Knots		
NYSE - Decimalisation (lower)	0.039	-0.005
NYSE - Decimalisation (upper)	0.078	-
NASDAQ - Decimalisation (lower)	1.157	-1.191
NASDAQ - Decimalisation (upper)	2.314	-
NYSE - Autoquote (lower)	0.090	-0.072
NYSE - Autoquote (upper)	0.180	-
NYSE - TSP Q (lower)	-0.053	0.066
NYSE - TSP Q (upper)	-0.106	-
NASDAQ - TSP Q (lower)	-0.053	0.078
NASDAQ - TSP Q (upper)	-0.106	-
Panel C. 15 Knots		
NYSE - Decimalisation (lower)	0.056	0.018
NYSE - Decimalisation (upper)	0.112	-
NASDAQ - Decimalisation (lower)	1.167	-1.285
NASDAQ - Decimalisation (upper)	2.334	-
NYSE - Autoquote (lower)	0.092	-0.085
NYSE - Autoquote (upper)	0.184	-
NYSE - TSP Q (lower)	-0.051	0.059
NYSE - TSP Q (upper)	-0.102	-
NASDAQ - TSP Q (lower)	-0.059	0.094
NASDAQ - TSP Q (upper)	-0.118	-

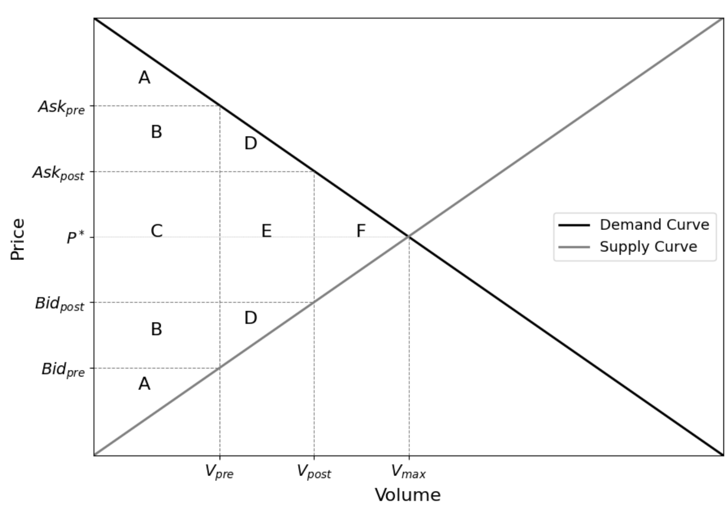
Figures

Figure 1
Gains from trade and cost of illiquidity
This figure graphically illustrates the gains from trades of a given stock for a given period and welfare cost of illiquidity. Panel A plots buyers' (demand curve) and sellers' (supply curve) private values and the corresponding volume from which we estimate gains from trades and welfare cost of illiquidity. Total buyer and seller surplus depends on the proportion of trades with a dealer-like counterparty (market maker). Buyer (seller) surplus ranges from half the upper (lower) 'T' triangle when a market maker is always counterparty to the full 'T' triangle when the market maker is never a counterparty. Market maker revenues ranges from zero (when market makers are never counterparty) to half the 'M' rectangle when market makers are always counterparty. Panel B illustrates gains from trade due to a reduction in bid-ask spread. Trader gains from trade originate from 'B' and 'D' while market makers lose revenue associated with 'B' but benefit from gain revenue associated with 'E'. Panel C transforms Panel B onto a plot with bid-ask spread on the y-axis and volume on the x-axis.

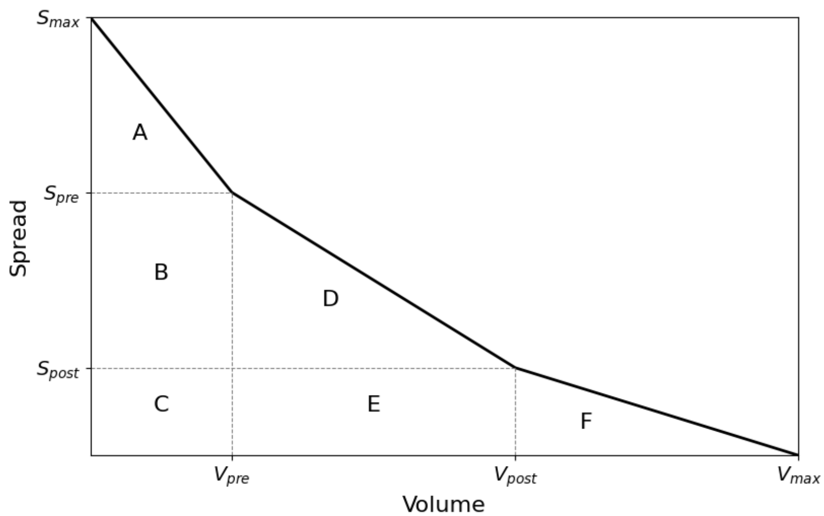
Panel A. Gains from trade and welfare cost of illiquidity



Panel B. Gains from trade due to change in liquidity

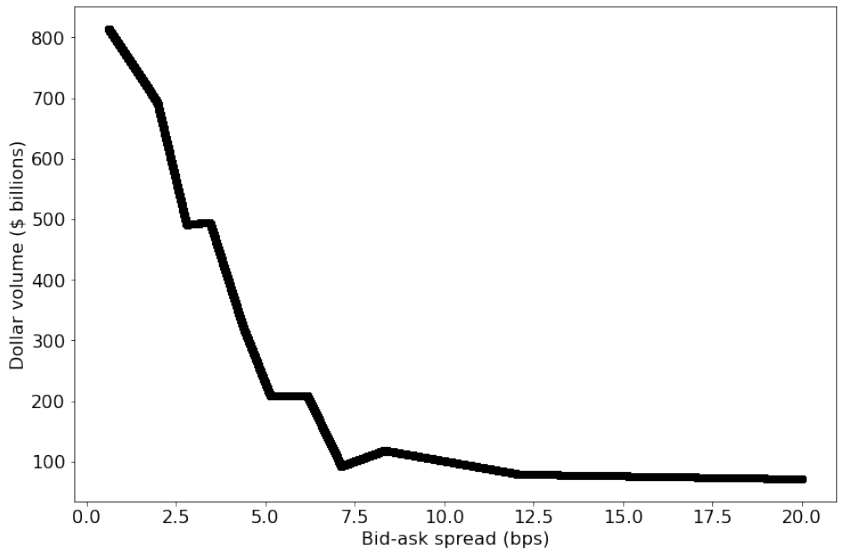


Panel C. Sensitivity of volume to bid-ask spreads



NVIDIA Demand Curve
This figure plots the bid-ask spread (in bps) against dollar volume ($ billions) for NVIDIA traded on the Nasdaq. Bid-ask spreads are calculated at a 5-minute frequency from 1 January 1996 to 31 December 2022. In Panel A, dollar volume is estimated using a spline regression with 9 knots corresponding to bid-ask spread deciles. In Panel B, dollar volume is estimated using 10 OLS regressions within each bid-ask spread decile.

Panel A. Spline regression



Panel B. OLS regression