The Hidden Cost of Zero-Commission*

Simon Jantschgi
University of Oxford†

First draft: February 2024
Current version: August 2024

Abstract

In today's financial landscape, traditional exchanges compete against online trading platforms. A critical point of competition centers around transaction costs. While traditional exchanges adhere to transparent transaction cost structures, many online trading platforms, under the guise of 'zero-commission trading', conceal transaction costs within their bid-ask spread. In this paper, I show that hidden transaction costs induce additional volatility in the form of price cycles in markets that would be stable if transparent transaction costs were charged. To compete with the profit opportunities from price cycles on platforms with hidden transaction costs, platforms with transparent costs must reduce them below the optimal monopolist level to attract traders. In this duopoly, I show that there is a market equilibrium: more risk-averse traders prefer transparent transaction costs, while less risk-averse traders choose hidden costs. Depending on the risk attitudes of traders, transparent transaction costs can be more or less efficient than hidden transaction costs. Finally, I show that the lack of commitment to transparent transaction costs can also lead to market failure, as patient traders may exploit price cycles through strategic market entry, creating a novel Coasian dynamic in a two-sided market.

Keywords: Market Design, Financial Markets, Transaction Costs
JEL Codes: D44, D47, G10, G12, G14, G23, L13

*I am grateful for comments and suggestions by Paul Klemperer, Scott Duke Kominers, Heinrich Nax, Bary Pradelski, Marek Pycia, Alex Teytelboym, participants of the Fall 2023 program 'Mathematics and Computer Science of Market and Mechanism Design' at the Simons Laufer Mathematical Sciences Institute, and the program committee and audience at ACM EC'24. This material is based upon work supported by the National Science Foundation under Grant No. DMS-1928930 and by the Alfred P. Sloan Foundation under grant G-2021-16778, while I was in residence at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Fall 2023 semester.

†Department of Economics and Nuffield College, OX1 3UQ, United Kingdom
E-Mail: simon.jantschgi@economics.ox.ac.uk

1 Introduction

In 2023, the global online trading platform market was valued at $43.04 billion USD with an expected compound annual growth rate of 8.6% between 2024 and 2032.[1] This growth highlights the increasing influence of online trading platforms such as Robinhood, E-Toro, E*TRADE, Charles Schwab, and WeBull in the financial landscape, positioning them as significant counterparts to traditional exchanges like the New York Stock Exchange (NYSE). A key competitive aspect between modern online trading platforms and traditional exchanges lies in the area of transaction costs. These costs are not only crucial to the functioning of financial markets, but they also play a significant role in influencing traders' decisions when selecting a marketplace to conduct their trading activities.

Traditional exchanges like the NYSE are known for their transparency and predictability in terms of transaction costs. This is partly due to the stringent regulations imposed by entities such as the Securities and Exchange Commission (SEC), which enforce strict standards of operation to ensure fair and orderly markets. Buy and sell prices are determined by prevailing market conditions, and transaction costs are transparent, often as percentages of transaction values or as flat rates per share, ensuring traders can readily discern their potential expenses.[2]

In contrast, modern online trading platforms promote the allure of zero-commission trading, enticing retail investors with the mirage of an ostensibly cost-free trading environment. However, a deeper look unveils a nuanced picture: transaction costs on these platforms have not vanished; they have merely transmuted. Zero-commission platforms predominantly monetize via two mechanisms: Payment for Order Flow (PFOF) — wherein platforms like Robinhood sell their order books to market makers such as Citadel Securities,[3] or by acting as market makers themselves, as seen with platforms like E-Toro. A market maker's modus operandi differs substantially from a traditional exchange. Instead of allowing prices to be determined by the natural interplay of market demand and supply, market makers set buy and sell prices according to their own discretion after observing their orderbook.[4] Market makers and online trading platforms offering zero-commission trading emphasize their critical role in ensuring liquidity, democratizing market access, and stabilizing markets during exogenous shocks by absorbing volatility.[5]

[1] See, e.g., the Online Trading Platform Global Market Report by Expert Market Research (2023).

[2] Tradition exchanges like the NYSE employ multifaceted fee structures that fluctuate based on diverse parameters, such as the types of participants (e.g., adding vs. removing liquidity) and the nature of the trade. For instance, a transaction that removes liquidity might be charged a fee of approximately $0.0030 per share. Alternatively, some trades are subject to fees as a percentage of the trade value, around 0.1%.

[3] As of the first quarter of 2023, around 70% of Robinhood's total net revenue of $441 million was generated through PFOF. Citadel Securities spent approximately $2.6 billion annually on PFOF across 2020 and 2021.

[4] U.S. Securities and Exchange Commission (2005) allows market makers some discretion in setting spreads to maintain a competitive, efficient market, permitting them to adjust spreads as market conditions change. However, it's important to distinguish these market makers from Designated Market Makers (DMMs) on traditional exchanges like the NYSE, who serve as liquidity providers by consistently posting transparent bids and asks.

[5] Citadel Securities emphasizes their role in ensuring quick and fair trading under all conditions, thereby generating market confidence (https://www.citadelsecurities.com/what-we-do/what-is-a-market-maker/).

However, the practices of market-making and PFOF have attracted significant regulatory attention.[6] Regulator's concerns are rooted in issues of transparency, and the overarching question of whether these mechanisms serve the best interests of retail investors or have the potential to disrupt the integrity of financial markets. Most of the scrutiny surrounding these practices hinges on the potential for conflicts in best price execution, particularly through the imposition of wider bid-ask spreads which can disadvantage the trader in pursuit of profits.

1.1 My Contribution

In this paper, I study how hidden transaction costs, attributable to market makers and zero-commission trading on modern online trading platform, compare to transparent transaction costs and ask the following question: *How does the (in)transparency of transaction costs impact a market, and what are the implications for both traders and the market platform?*

I consider a dynamic market environment, where traders arrive on a market platform at a constant rate. They can either be sellers, who wish to sell, or buyers, looking to purchase a homogeneous commodity. Upon entry, traders place limit orders, detailing the maximal price at which they are willing to purchase the commodity or the minimal price at which they are willing to sell the commodity. At fixed time steps, the market platform observes the orderbook, determines a buy and a sell price for the commodity, and executes subsets of buy and sell orders at those prices. The difference between the buy and sell price, known as the bid-ask spread, is the platform's revenue. The market platform sets the bid-ask in one of two ways:

1. *Hidden transaction costs (e.g. zero-commission models).* The platform sets buy and sell prices at their own discretion to maximize revenue after observing the orderbook.
2. *Transparent transaction costs (e.g. fees on traditional exchanges).* The platform sets buy and sell prices as ex-ante communicated functions of the market-clearing price.

Traders, whose order was executed, exit the market, while traders, whose order was not executed, remain for the next trading round with some probability or cancel their order.

The first contribution of this paper is to demonstrate that while transparent and hidden transaction costs may produce equivalent market outcomes in a static environment — yielding similar revenue for the platform and surplus for traders — this equivalence drastically breaks down in a dynamic market environment. Hidden transaction costs introduce volatility by creating price cycles in markets that would otherwise remain stable under transparent transaction costs. This finding directly challenges the conventional view that market makers and online trading platforms reduce volatility. The volatility arises because market makers

[6] The SEC in the U.S. has scrutinized PFOF, addressing concerns in their Regulation Best Execution proposal Securities and Exchange Commission (2022). In Europe, PFOF faces bans, with the UK's FCA leading since 2012 Financial Conduct Authority (2012), and ESMA finding it incompatible with MiFID II, indicating stricter EU standards European Securities and Markets Authority (2021). Furthermore, market makers are rigorously regulated by the SEC's Division of Trading and Markets to ensure fair trading practices and maintain market integrity.

can exploit accumulated excess demand and supply by adjusting the spread. As profitable orders accumulate unexecuted due to spread constraints, the market platform can profit by temporarily narrowing the spread to increase trading volume. Once the excess demand or supply is cleared, the platform reverts to a wider spread, creating cyclical price movements. In contrast, a platform with transparent transaction costs cannot engage in such adjustments. Because these platforms commit upfront to a fixed cost structure, they cannot manipulate the spread to clear accumulated orders. The key reason is that the accumulated unexecuted orders do not influence the market-clearing price under transparent costs. As a result, the bid-ask spread remains constant over time, leading to a more stable and predictable market environment.

The second contribution is to show how the additional volatility on market platforms with hidden transaction costs affects its competition against a market platform with transparent transaction costs. I establish the existence of a market equilibrium, where transparent transaction costs must be charged below the optimal monopolist level to compete against the additional profit opportunities from price cycles. The optimal transparent transaction cost depends on the risk attitude of traders. Higher levels of risk aversion allow a market platform to charge transparent transaction costs at a higher rate. Surprisingly, there is a cross-over effect, when it comes to average market efficiency. For high levels of risk aversion, transparent transaction costs are on average less efficient than hidden transaction costs, but this reverses for low levels of risk aversion, where transparent transaction costs are more efficient. If traders have different risk attitudes, in equilibrium, there is a natural market segmentation: Traders with higher risk aversion favor the reliability of transparent transaction costs, while traders with lower risk aversion favor the additional profit opportunities of hidden transaction costs.

The third contribution of this paper is to study how the sophistication of traders affects market dynamics. In markets with transparent transaction costs, the spread remains stable and predictable over time, which means traders cannot influence their deals, making price-taking the optimal strategy. However, in markets with hidden transaction costs, patient traders can exploit predictable price cycles through a sophisticated strategy of market entry, potentially bringing the market to failure. This market failure represents a novel Coasian dynamic in a two-sided environment, where the platform's inability to commit to transparent transaction costs is the root cause of instability. Unlike classical Coasian scenarios where traders exert market power by waiting for favorable prices, here, prices are not announced in advance. Thus, patient traders cannot force the spread to narrow simply by waiting. However, if patient traders can anticipate the structure of price cycles, they can place aggressive orders that only execute when market conditions are favorable. If all traders adopt this strategy, the market's orderbook drastically narrows, causing price cycles to evolve with a smaller baseline spread. Anticipating this, iterative reasoning leads to best response strategies where the revealed demand and supply only support spreads that are unprofitable for the platform if it has to cover fixed costs. As a result, the market fails. This outcome theoretically supports the folk-wisdom that zero-commission market platforms with hidden transaction costs must attract unsophisticated traders to remain profitable.

1.2 Related literature

The importance of transaction costs in markets has been recognized at least since Coase (1960). Demsetz (1968) pioneered the analysis of the impact of such costs on financial markets.

Transparent transaction costs are well-studied in continuous-time models (representing stock exchanges in business hours) and discrete-time batch auctions (similar to off-hours clearing). Their impact is summarized as follows: First, higher transaction costs lead to a broader bid-ask spread, reducing trading volume, see, e.g., Barclay et al. (1998) and Noussair et al. (1998). Second, competition between exchanges generally decreases transaction costs, see Cantillon and Yin (2011) for a survey. Third, the break-down of fees between liquidity demanders and suppliers matters (Colliard and Foucault 2012, Foucault et al. 2013, Malinova and Park 2015). For discrete time models, transparent transaction costs have received attention in the literature on mechanism design, analyzing their impact on trader's incentives and market efficiency (Tatur 2005, Chen and Zhang 2020, Jantschgi et al. 2023). I add to this literature by providing a novel model to study transparent transaction costs, and show that they lead to stable outcomes in dynamic markets.

Hidden transaction costs emerge when market makers set bid-ask spreads at their discretion. The market microstructure literature typically views market makers as providers of immediacy, with the spread as their service fee for providing and holding risky assets, see O'Hara (1998) or Madhavan (2000). The main thrust of this literature is concerned with how the spread is affected by traders with different levels of asymmetric information (Glosten and Milgrom 1985, Kyle 1985) and risk attitudes of market makers (O'Hara and Oldfield 1986) in the face of uncertainty. Their key finding is that uncertainty, insider trading risks, and varying risk aversion levels can result in a dynamic spread. I extend this literature by showing how the accumulation of excess demand and supply prompts market makers to modify the bid-ask spread, even without uncertainty, risk, or insider information, in a framework where the asset's value is dictated entirely by current demand and supply. This insight starkly contrasts with the classical belief that market makers stabilize markets and volatility is only caused by informational asymmetries and external shocks.

It has been well-established that market makers, zero commission models, and hidden transaction costs are linked to PFOF (Chordia and Subrahmanyam 1995, Kandel and Marx 1999, Battalio and Holden 2001, Parlour and Rajan 2003). The main thrust of the literature on PFOF finds that it leads to wider bid-ask spreads, justifying regulatory concerns regarding best execution guarantees. More recent work includes Battalio and Loughran (2008), Anand et al. (2016), Battalio et al. (2016), Ernst and Spatt (2022) for more recent work. This paper adds to this literature by providing a novel theoretical model to study zero-commission models and by uncovering new layers of potential conflicts like additional volatility, strategic incentives, and profit opportunities. Moreover, by studying platform competition, I provide theoretical justification for the empirical finding that hidden transaction costs can lead to wider bid-ask spreads than transparent transaction costs.

High volatility and cyclical price patterns are well-documented in various economic contexts. Most relevant to this paper is the cyclical pricing observed for a durable goods monopolist, where the arrival of new traders and the accumulation of excess demand lead to additional volatility, as demonstrated by Conlisk et al. (1984) and Sobel (1991). In financial markets, significant price jumps are frequently linked to major news revelations following information aggregation failures, often due to irrational behaviors like informational cascades (Bikhchandani et al. 1992) and herding (Banerjee 1992, Shiller 1995). Some models can generate big price movements without substantial news, based on strategic interactions of rational traders (Bulow and Klemperer 1994) and failure of information aggregation due to transaction costs (Lee 1998). In gasoline retail markets, asymmetric price cycles with small price decreases and large price increases are regularly observed (Eckert 2002, Noel 2007, Wang 2009). These so-called Edgeworth cycles, often linked to tacit collusion, can be explained by oligopolistic price competition (Maskin and Tirole 1988, Noel 2008). In particular, my model is similar to Nisan (2023)'s work on cyclical price patterns on blockchains driven by excess demand. Complementary, this paper explains cyclical price patterns on two-sided market platforms, even in the absence of irrational behavior or competition, as a result of hidden transaction costs and accumulation of excess demand and supply, similar to a durable good monopolist, who sets transaction costs instead of prices.

My paper connects to the dynamic revenue management literature (van Ryzin and Talluri (2005),McAfee and Wiseman (2008), Hörner and Samuelson (2011), Board and Skrzypacz (2016)), and to studies exploring monopolists' decisions between price discrimination and fixed pricing in the presence of strategic buyers, particularly in relation to the Coase Conjecture (Coase (1972), Stokey (1981), Gul et al. (1986), Board and Pycia (2014), Brzustowski et al. (2023)). The key contribution of this paper is uncovering a novel Coasian dynamic within a two-sided market, where the monopolist is not pricing goods but rather setting transaction costs. Unlike traditional Coasian scenarios, where strategic buyers exert power by timing their entry, in this environment, prices are not publicly announced. Instead, traders exercise their market power by strategically submitting aggressive orders that only execute under favorable market conditions.

Finally, by examining competition among financial platforms with varying transaction cost models, I add to the extensive body of research on platform competition, c.f., Rochet and Tirole (2003) and Armstrong (2006). This paper specifically contributes to the discourse on fee competition, see, e.g., Weyl (2010), Tan and Zhou (2021), and Teh et al. (2023).

2 The Model

2.1 The Static Market

The Traders. I consider a market in which a continuum of traders play one of two roles: sellers sell and buyers buy units of a homogeneous commodity. $\mathcal{B}$ denotes the set of buyers and $\mathcal{S}$ denotes the set of sellers, which are closed intervals in $\mathbb{R}$.[7] Every trader $i \in \mathcal{B} \cup \mathcal{S}$ submits a *limit order* $o_i = (q_i, v_i)$ to a market platform, consisting of the quantity q_i that they want to trade, and a value v_i. Without loss of generality, I assume that for every trader i, $v_i \in [1, 2]$. For a buyer b, the value v_b is called a *bid* and represents the maximum price at which they are willing to purchase one unit of the good. For a seller s, this value v_s is called an *ask* and represents the minimum price at which they are willing to sell one unit the good. Let $q_B : \mathcal{B} \to [0, 1]$ and $q_S : \mathcal{S} \to [0, 1]$ denote Borel-functions assigning each buyer and seller their preferred quantity of trade. Similarly, let $v_B : \mathcal{B} \to [1, 2]$ and $v_S : \mathcal{S} \to [1, 2]$ denote Borel-functions assigning each trader their value. Let μ_B and μ_S be two absolutely continuous measures on $\mathcal{B}$ and $\mathcal{S}$ with densities $q_B(\cdot)$ and $q_S(\cdot)$. μ_B and μ_S describe the distribution of traders, with their mass being equal to their preferred quantity of trade. Let μ_B^t and μ_S^t denote the push-forward measures of μ_B and μ_S on $[1, 2]$ via the functions v_B and v_S. μ_B^t and μ_S^t describe the distribution of values on $[1, 2]$, taking their size into account.

Demand and Supply. Buy and sell orders are aggregated to *demand* and *supply functions*, which specify the mass of traders, who, given their orders, are willing to trade at price P.[8]

$$\begin{aligned} D(P) &= \mu_B(\{b \in \mathcal{B} : v_b \geq P\}) = \mu_B^t([P, 2]) \quad (\textit{Demand Function}) \\ S(P) &= \mu_S(\{s \in \mathcal{S} : v_s \leq P\}) = \mu_S^t([1, P]) \quad (\textit{Supply Function}) \end{aligned}$$

Assumption (Analytical Properties of Demand and Supply). Demand and supply are $\mathcal{C}^1$-functions, with demand being strictly decreasing on $[1, 2]$ and supply being strictly increasing on $[1, 2]$. Their derivatives are strictly bounded from above and below on $[1, 2]$. The total mass of orders is one for each market side.

The unique *market-clearing price* P_{eq} equates demand and supply, that is, $D(P_{eq}) = S(P_{eq})$.

The Market Platform. A market platform facilitates trade as follows: It sets a buy price P_b and a sell price P_s and execute subsets $\mathcal{B}^* \subset \mathcal{B}$ and $\mathcal{S}^* \subset \mathcal{S}$ of buy and sell orders. Buyers with executed order pay the buy price P_b per unit, and sellers with executed order receive the sell price P_s per unit. A key constraint is *trade-balance*, that is, $\mu_B(\mathcal{B}^*) = \mu_S(\mathcal{S}^*)$. The gap between buy and sell price, called the *bid-ask spread*, is $\sigma = P_b - P_s \geq 0$. Market platforms determine the terms of trade according to one of two *transaction costs (TCs)*:

[7] Studying markets with continuum of traders dates back to Aumann (1964).

[8] Jantschgi et al. (2022) demonstrate that in finite markets with n traders on both sides, demand and supply step functions approximate continuous curves at a rate of $\mathcal{O}(n^{-1/2})$. Considering the usual depth of order books, a trader continuum is thus an effective approximation.

- *Transparent TCs.* The platform executes all buy orders above and all sell orders below the market clearing price: $\mathcal{B}^* = \{b \in \mathcal{B} : t_b \geq P_{eq}\}$ and $\mathcal{S}^* = \{s \in \mathcal{S} : t_s \leq P_{eq}\}$. Buy and sell prices are set as continuous and strictly increasing functions of P_{eq}, which the traders know ex-ante: $P_b = F_B(P_{eq}) \geq P_{eq}$ and $P_s = F_S(P_{eq}) \leq P_{eq}$.[9] For example, a *price fee* $p \in (0, 1)$ with $F_B(P) = (1+p)P$ and $F_S(P) = (1-p)P$.

- *Hidden TCs.* The platform sets buy and sell price at their own discretion after observing demand and supply. All buy orders with bid above the buy price and sell orders with ask below the sell price are executed, that is, $\mathcal{B}^* = \{b \in \mathcal{B} : t_b \geq P_b\}$ and $\mathcal{S}^* = \{s \in \mathcal{S} : t_s \leq P_s\}$. Trade-balance imposes $D(P_b) = S(P_s)$.

I use the term *transparent* for transaction costs directly tied to the market clearing price. When traders know the market environment and equilibrium price, they can fully discern their transaction costs and the resulting buy and sell prices. In contrast, *hidden* transaction costs occur when the platform sets buy and sell prices at its discretion, making it impossible to separate the market price from additional fees ex-ante. As noted in the Introduction, transparent costs are common on major stock exchanges, while hidden costs often appear on zero-commission platforms that sell order flow to third-party market makers. In this model, the market platform combines the roles of a zero-commission broker collecting limit orders and the market maker executing the trades.

Trader's Behavior. A trader's true value is such that they are indifferent to trading, if it coincides with their total payment. In order to not make losses, traders must guarantee *ex-post individual rationality*: buyers should trade only if the buy price is below their value, and sellers should trade only if the sell price exceeds theirs. The *net value* v_b^{net} of a buyer b is the largest bid that guarantees ex-post individual rationality. Similarly, the *net value* v_s^{net} of a seller s is the smallest ask that guarantees ex-post individual rationality.

- *Transparent TCs.* Net values are different from true values. The set of executed orders depends on the market-clearing price, with transaction costs added ex-post. If traders report their true value, they risk trades where the market-clearing price matches their value, leading to a loss due to transaction costs. For a buyer b, the net value v_b^{net} is $v_b^{net} = F_B^{-1}(v_b)$. For a seller, the net value v_s^{net} is $v_s^{net} = F_S^{-1}(v_s)$.[10] Thus, net values are continuous and strictly increasing in the true value. For a price fee p, net values are $v_b^{net} = \frac{v_b}{1+p_b}$ and $v_s^{net} = \frac{v_s}{1-p_s}$.

- *Hidden TCs.* Net values are equal to true values. Reporting the true value is ex-post individually rational. That is, because a buyer is involved in trade, if and only if their bid is above the buy price, and a seller is involved in trade, if and only if their ask is below the sell price.

[9] Strict monotonicity ensures that buyers prefer lower market-clearing prices, while sellers prefer higher ones. Equivalently: In addition to P_{eq}, the platform charges transaction costs as functions of P_{eq}, $\Phi_B(P_{eq})$ and $\Phi_S(P_{eq})$. The buy price is $P_b = P_{eq} + \Phi_B(P_{eq})$ and the sell price is $P_s = P_{eq} - \Phi_S(P_{eq})$.

[10] This follows from the necessary and sufficient equations $v_b - F_B(v_b^{net}) = 0$ and $F_S(v_s^{net}) - v_s = 0$. A formal derivation of net values can be found in Jantschgi et al. (2023).

Reporting the net value is the least aggressive bid that guarantees ex-post individual rationality. I call this behaviour *truthfulness* or *price-taking*. Let $D^0(\cdot)$ and $S^0(\cdot)$ denote the *net demand and supply*, adjusting true demand and supply for net values. Given that net values are continuous and strictly increasing, net demand and supply inherit the analytical properties from true demand and supply.

Assumption (Price-Taking Behavior). Until Section 6, I assume that upon arrival on the platform, traders act as price-takers and do not strategically misrepresent their value.

In a static market, price-taking is the optimal strategic.[11] However, in the dynamic market, it will turn out that this behaviour can be thought of as *myopic*, as traders, who properly understand the market evolution outlined in Sections 3 and 4, can, in some cases, benefit from strategically misrepresenting their bid, see Section 6.

Market Metrics. *Trade volume* is the mass of all active buyers and sellers. For an active trader, *welfare* is the difference between their valuation and the price, scaled by the order size. *Total Welfare* is given by the welfare of all active traders. *Revenue* is the bid-ask spread, times the trading volume. *Realized gains of trade* are the sum of welfare and revenue. The *gross gains of trade* correspond to the maximum possible gains, which occur when the spread is zero and trade happens at the equilibrium price. The *loss* measures, how much gains of trade are lost due to a strictly positive spread. The *market efficiency* measures as a percentage, how many gains of trade are realized. More formally:

$$Q = \mu_B(\mathcal{B}^*) = \mu_S(\mathcal{S}^*) \qquad \text{(Trade Volume)}$$

$$W = \int_{\mathcal{B}^*} (v_b - P_b)\, d\mu_B(b) + \int_{\mathcal{S}^*} (P_s - v_s)\, d\mu_S(s) \qquad \text{(Welfare)}$$

$$R = \int_{\mathcal{B}^*} P_b\, d\mu_B(b) - \int_{\mathcal{S}^*} P_s\, d\mu_S(s) = (P_b - P_s) \cdot Q \qquad \text{(Revenue)}$$

$$G^{real} = \int_{\mathcal{B}^*} v_b\, d\mu_B(b) - \int_{\mathcal{S}^*} v_s\, d\mu_S(s) \qquad \text{(Realized Gains of Trade)}$$

$$G^{gross} = \int_{\mathcal{B}^{eq}} v_b\, d\mu_B(b) - \int_{\mathcal{S}^{eq}} v_s\, d\mu_S(s) \qquad \text{(Gross Gains of Trade)}$$

$$L = G^{gross} - G^{real} \qquad \text{(Loss)}$$

$$E = \frac{G^{real}}{G^{gross}} \qquad \text{(Efficiency)}$$

The gross gains of trade are thus equal to the sum of traders' welfare, revenue and loss, that is $G^{gross} = W + R + L$. We identify *market performance* with the triple (W, R, L).

[11] Jantschgi et al. (2023) find that in finite markets, strategic behavior in the presence of transaction costs converges to price-taking at a rate $\mathcal{O}(\frac{1}{n})$. Therefore, under the large market assumption, price-taking accurately reflects optimal market behavior.

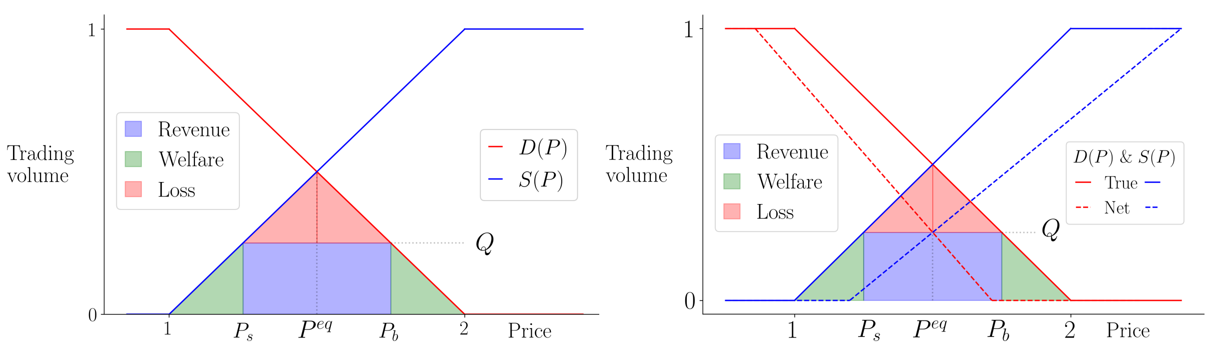
Static Market Equivalence

In a static market, it turns out that hidden and transparent transaction costs are equivalent:

Proposition 2.1 (Static equivalence)**.** *In a static market, the same set of market performances can be achieved with transparent transaction costs and hidden transaction costs.*

Intuition. For hidden transaction costs, the market performance is fully specified by the rectangle with height equal to the trading volume and width equal to the chosen spread that fits under the true demand and supply curve, see left-hand side of Figure 1. For transparent transaction costs, the market performance is also fully specified by the trading volume, that is, the intersection of revealed demand and supply. The market performance is then equal to the largest rectangle with height equal to that trading volume that fits under the true demand and supply curve, see right-hand side of Figure 1. For any height of that rectangle, transparent transaction costs can be scaled such that they intersect exactly at that height. The formal of this last claim can be found in Jantschgi et al. (2023, Proposition12).

Example 2.2 (Static market with linear demand and supply)**.** Suppose that true values of traders are such that demand is $D(P) = 2 - P$ and supply is $S(P) = P - 1$ on $[1, 2]$. If the market platform sets the spread at their own discretion, traders report their true value. Thus, revealed demand and supply are $D^0(P) = 2 - P$ and $S^0(P) = P - 1$ on $[1, 2]$. If the market platform charges a price fee p, that is, $P_b = (1 + p)P^{eq}$ and $P_s = (1 - p)P^{eq}$, traders report their net value, that is, $t_b^{net} = {t_b}/{(1+p)}$ for buyers and $t_s^{net} = {t_s}/{(1-p)}$ for sellers. Thus, revealed demand and supply are $D^0(P) = 2 - (1 + p)P$ and $S^0(P) = (1 - p)P - 1$. Figure 1 compares revealed demand and supply, as well as the market performance, for the two transaction cost models.



((a)) **Choosing the spread** $\sigma = P_b - P_s$ ((b)) **Choosing the price fee** p

Figure 1: **Market performance in a static market.** This figure shows the market performance for hidden transaction costs (Left) for a transparent price fee (Right). In both cases, the market performance is fully specified by the trading volume: The revenue (blue) is equal to the largest rectangle with height equal to the trading volume that fits under true demand and supply. The trader's surplus (green) is the area to the left and right of the revenue rectangle. The loss (red) is the area above the revenue rectangle.

The following table shows analytical expressions for the market performances:

	Market Price	Buy price	Sell price	Volume	Revenue	Welfare	Loss
Spread σ	$\frac{3}{2}$	$\frac{3+\sigma}{2}$	$\frac{3-\sigma}{2}$	$\frac{1-\sigma}{2}$	$\frac{\sigma-\sigma^2}{2}$	$\frac{1-2\sigma+\sigma^2}{4}$	$\frac{\sigma^2}{4}$
Price fee p	$\frac{3}{2}$	$\frac{3(1+p)}{2}$	$\frac{3(1-p)}{2}$	$\frac{1-3p}{2}$	$\frac{3p-9p^2}{2}$	$\frac{1-6p+9p^2}{4}$	$\frac{9p^2}{4}$

In line with Proposition 2.1, the two transaction cost structures are equivalent. That is, for $\sigma = 3p$, the market performances coincide. For linear demand and supply, revenue is maximized, when the spread is set as $\sigma = 0.5$, or equivalently, if the price fee is set as $p = \frac{1}{6}$.

2.2 The Dynamic Market

Next, I consider a dynamic extension of the static market environment.

Arrival of Traders. At discrete time steps $t = 1, 2, \ldots$, new buyers and sellers enter the market platform. Real-world markets often face exogenous shocks, causing dynamic shifts in demand and supply. One key result of this paper is that transaction costs can potentially introduce volatility into an otherwise stable market. To minimize the impact of modeling choices on volatility, I assume the following for the main part of the paper:

Assumption (Homogeneous Arrival). The arrival of new traders is *homogeneous*, that is, the value distributions of incoming buyers and sellers are constant over time. Let $D^0(P)$ and $S^0(P)$ be the *baseline demand* and *supply*, already accounting for net values. Incoming demand and supply satisfy the analytical properties imposed in Section 2.1. Moreover, I assume that for $D^0(\cdot)$ and $S^0(\cdot)$, there exists a unique revenue-maximizing spread σ^0.

This provides the foundation for analyzing market dynamics without introducing volatility through the model itself. The uniqueness of σ^0 will ensure that there are no selection and tie-breaking issues in the baseline market. This assumption can be relaxed by allowing for exogenous shocks to demand and supply. The main results of the paper qualitatively extend.[12]

Clearing Events. At each time step t, the market platform observes the current market condition, that is, the revealed demand and supply $D^t(P)$ and $S^t(P)$ of all traders currently in the market. As in Section 2.1, the platform sets buy and sell prices using hidden or transparent transaction costs, and executes subsets of active limit orders.

Thus, the dynamic market model is that of *frequent batch auctions*, where multiple orders are executed simultaneously. Although traditional financial exchanges usually operate in continuous time, they use batch auctions to clear accumulated off-hours offers, determining the next day's opening price. Budish et al. (2015) suggested that frequent batch auctions could mitigate the '*High Frequency Arms Race*' by reducing mechanical arbitrage rents. Moreover,

[12] Detailed simulations of this claim will be available in the near future in an Online Appendix.

market makers also frequently bundle client's orders. Hence, frequent batch auctions provide a realistic and unified mechanism to understand the impact of different transaction costs.

It is important to note that this model does not address the frequency at which frequent batch auctions should occur; instead, it assumes a given frequency with associated incoming demand and supply as exogenously determined. The trade-off between longer trading periods, which increase market thickness, and shorter trading windows, which reduce waiting times, lies beyond the scope of this paper, and is well-studied (without transaction costs), see, e.g. Mendelson (1982), Loertscher et al. (2022). However, this trade-off is particularly relevant in the context of contracting Payment for Order Flow and for designing optimal batch auctions with transaction costs, and is thus an interesting avenue for future research.

Departure of Traders. Traders, whose order was executed in round t, leave the market. Traders with unfulfilled orders remain in the market for round $t+1$ with some probability. Let $Z_D^t(P)$ denote the built-up demand corresponding to unfulfilled buy orders after round t. Similarly, let $Z_S^t(P)$ denote the built-up supply after round t. More formally:

$$Z_D^t(P) = \left(D^t(P) - Q^t\right) \mathbf{1}_{\{P \leq P_{eq}^t\}} \text{ and } Z_S^t(P) = \left(S^t(P) - Q^t\right) \mathbf{1}_{\{P \geq P_{eq}^t\}} \quad (\textit{Transparent TCs})$$

$$Z_D^t(P) = \left(D^t(P) - Q^t\right) \mathbf{1}_{\{P \leq P_b^t\}} \text{ and } Z_S^t(P) = \left(S^t(P) - Q^t\right) \mathbf{1}_{\{P \geq P_s^t\}} \quad (\textit{Hidden TCs})$$

Assumption (Memoryless Departure). The departure of traders with unfulfilled orders is *memoryless*: Limit orders, which are not executed in round t, remain in the market for round $t+1$ with probability ϵ, or are deleted with probability $1-\epsilon$. Thus, at round $t+1$, demand and supply are

$$D^{t+1}(P) = \epsilon \cdot Z_D^t(P) + D^0(P) \text{ and } S^{t+1}(P) = \epsilon \cdot Z_S^t(P) + S^0(P) \quad (\textit{Market Evolution})$$

For an illustration of the demand and supply dynamics in the first trading rounds, please refer to Figure 2 (b) in Section 3 for hidden transaction costs, and to Figure 4 (b) in Section 4 for transparent transaction costs.

In financial markets, unexecuted limit orders remain in the order book, awaiting price movements for execution.[13] Market makers hold orders until execution or client cancellation, not deleting them arbitrarily. Therefore, the assumptions on departure and market evolution, where orders remain active until fulfilled or canceled, align well with real-world practices. The assumption of equal departure rates for buyers and sellers is not crucial to this analysis. The results for transparent transaction costs remain unchanged, while the results for hidden transaction costs are qualitatively extended. The main technical challenge arises from the asymmetric buildup of excess demand and supply, which skews the market over time.[14]

[13] Hollifield et al. (2006) found that around 70% of welfare loss was due to non-execution of limit orders, often because of transaction costs.

[14] An Online Appendix with detailed simulations will be available soon.

Long-Run Performance. The superscript t denotes market metrics in round t. Let $\Sigma = (\sigma^t)_{t\geq 1}$ denote a sequence of spreads. Define the *long-run averages* as

$$W^\infty(\Sigma) = \lim_{T\to\infty} \frac{\sum_{t=1}^T W^t(\sigma^t)}{T} \qquad (\textit{Average Welfare})$$

$$R^\infty(\Sigma) = \lim_{T\to\infty} \frac{\sum_{t=1}^T R^t(\sigma^t)}{T} \qquad (\textit{Average Revenue})$$

$$L^\infty(\Sigma) = \lim_{T\to\infty} \frac{\sum_{t=1}^T L^t(\sigma^t)}{T} \qquad (\textit{Average Loss})$$

$$E^\infty(\Sigma) = \lim_{T\to\infty} \frac{\sum_{t=1}^T E^t(\sigma^t)}{T} \qquad (\textit{Average Efficiency})$$

Note that for average efficiency, if G_{gross} represents the gross gains of trade for incoming value distributions $D^0(\cdot)$ and $S^0(\cdot)$, then $E^\infty(\sigma) = \lim_{T\to\infty} \frac{\sum_{t=1}^T G^t_{real}(\sigma^t)}{T\cdot G_{gross}}$. The definitions are equivalent as trade timing is irrelevant for the long-run average. Each round adds G^{gross}, so $T \cdot G^{gross}$ measures potential trade over T rounds, and $\sum_{t=1}^T G^t_{real}$ measures realized gains.

3 Hidden TCs lead to Predictable Price Cycles

In the following section, I analyze the evolution of the bid-ask spread, if the platform charges hidden transaction costs. For the main part of the paper, I make the following assumption:

Assumption (Myopic Revenue Maximization). The platform charging hidden transaction costs sets the spread to myopically maximize per-round revenue, i.e., $\sigma = \arg\max_{\sigma\in[0,1]} R(\sigma)$. In case of ties, the platform opts for the smaller spread with a larger trading volume.

This assumption facilitates a more straightforward and intuitive analysis of the primary economic driving forces. While myopic revenue maximization may be reasonable due to market volatility and the unpredictability of long-term forecasts — especially in high-frequency trading, where operations exploit quick-changing price variations — it's also natural to study the maximization of long-run average revenue. In Appendix A.2, I extend the analysis to a platform selecting a sequence of spreads to maximize average revenue. The main qualitative insights of Theorem 3.1 and Theorem 3.2 below extend, as demonstrated in Theorem A.2.

Existence of Price Cycles. In a first step, I show that the sequence of realized spreads is not constant over time, even though the arrival and departure of traders is homogeneous. Instead, there emerge price cycles. Recall that for the incoming baseline demand and supply, I assumed that the revenue-maximizing spread is unique and denoted by σ^0. This *baseline spread* σ^0 will play an important role, as the starting point of the price cycles.

Theorem 3.1 (Hidden TCs $\Rightarrow$ Price Cycles.). *If a market platform charges hidden transaction costs, the market-clearing price is constant over time. However, the bid-ask spread evolves as follows:*

- ***Monotonicity.*** *For all* $t > 1$, $\sigma^t = \sigma^0$ *or* $\sigma^t < \sigma^{t-1}$.
- ***Upper*** & ***lower bound.*** *There exists a spread* $\sigma^* > 0$, *such that for all* $t > 1$, $\sigma^* \leq \sigma^t \leq \sigma^0$.
- ***Cycles.*** *There exist infinitely many* t *with* $\sigma^t = \sigma^0$. *If the departure rate* $1 - \epsilon$ *is sufficiently low, there exist infinitely many* t' *with* $\sigma^t \neq \sigma^0$.

This result may seem surprising at first glance. The market environment is stable, as traders arrive and depart at a constant rate, leading to a stable market-clearing price. Why would it be optimal for a platform to adjust the spread over time?

Intuition. After some time, the platform can profit from *built-up excess demand and supply*. In the second round, there is built-up demand on the interval $[P_b^0, 2]$ and built-up supply only on the interval $[1, P_s^0]$ from orders, which were not executed in the first round, and which remain active. If the departure rate is sufficiently low, after some trading rounds at the baseline spread, enough excess demand and supply has accumulated, such that it becomes profitable to lower the spread and clear more trades. However, once the market is cleared at a certain spread, all the excess demand and supply for this spread is gone. Thus, if the market were to clear again at the same spread. It would clear only baseline demand and supply, for which the baseline spread σ^0 is optimal. Hence, after deviating from the baseline spread, the spread must keep on strictly decreasing or jump back up to the baseline spread. However, due to the non-zero departure rate, excess demand and supply does not build up forever. Thus, at some point, most profitable excess demand and supply has been cleared. At this point, it is optimal for the platform to revert back to a wider spread that is optimal for the baseline demand and supply. Thus, a next cycle starts. The formal proof of Theorem 3.1 is relegated to Appendix B.2.

Predictability of Price Cycles. Next, I demonstrate that the sequence of the bid-ask spread is not just unstable; it follows a consistent and predictable pattern. To describe this pattern, I introduce the following notation: Let $\Sigma^{[T_1,T_2]} = (\sigma^t)_{T_1 \leq T \leq T_2}$ denote the sequence of realized spreads in the time interval $[T_1, T_2]$. A key focus will be on identifying points in time when a new lowest-ever spread occurs, followed immediately by a spread that returns to the baseline level. I define the *n'th price cycle* C_n as

$$C_n = \Sigma^{[t_{n-1}+1,t_n]} \text{ for } t_0 = 0 \wedge t_n = \inf\{t \geq t_{n-1} : \sigma^t \leq \sigma^{t_{n-1}} \wedge \sigma^{t+1} = \sigma^0\}. \quad \textit{(Price Cycles)}$$

The next theorem shows price cycles are finite and follow a recurrent pattern.

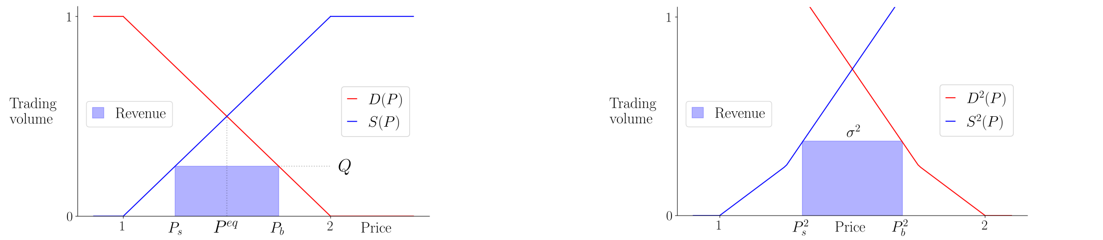
Theorem 3.2 (Predictability of price cycles)**.** *Consider a market platform with hidden transaction costs. The n^{th} price cycle is finite.* $\forall t \in [t_{n-1}, t_n]$, *if* $\sigma^t \geq \sigma^{t_{n-1}}$, *then* $\sigma^t = \sigma^{t-t_{n-1}}$.

Let $\bigcup_{i=1}^{n-1} C_i$ denote the concatenation of the first $n-1$ price cycle. Theorem 3.2 implies that n'th price cycle C_n repeats $\bigcup_{i=1}^{n-1} C_i$, until it is broken by a new global minimum. This theorem has multiple interesting implications. First, each spread, that is realized for the first time, must be a new global minimum. Second, the endpoints of each cycle, that is, σ^{t_n}, a global minimum. This sequence is monotone. Thus, the lower bound in Theorem 3.1 implies that the sequence of the minimal spread is converging over time. Moreover, the theorem implies that sequence of realized bid-ask spreads is countable, possibly finite. Excess demand and supply builds up continuously over time, changing the market environment. So why does the same pattern of prices emerge consistently?

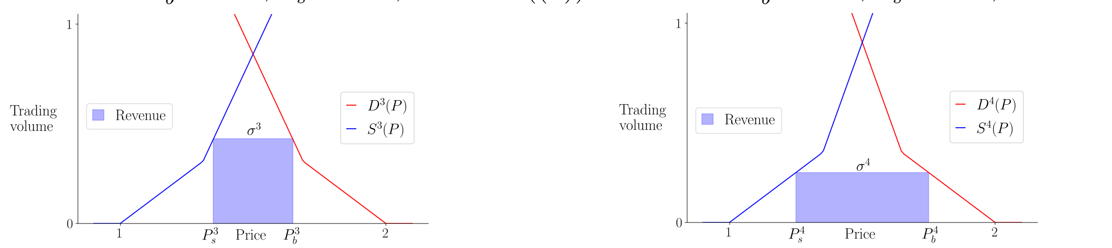
Intuition. At the end of every price cycle, the market was cleared at the global minimum spread. Thus, the only way to profit from the excess demand and supply from past price cycles is to clear the market at a new global minimum. For any wider spread, excess demand and supply that is cleared in the n'th cycle is the one that built up during that particular cycle. Given that the arrival process is homogeneous and the departure process is memoryless, excess demand and supply builds up in the exact same way as in the $n-1$'st price cycle. Thus, it is optimal to clear the market in the same way, unless the price cycle is broken by clearing the market at a smaller spread than ever before. The proof of Theorem 3.2 is relegated to Appendix B.3.

Example 3.3 (Hidden TCs in a market with linear demand and supply)**.** Suppose that the incoming distribution of demand and supply is linear, see Example 2.2, and that traders with unexecuted orders depart at a rate $1 - \epsilon = 0.05$. That is, before the second trading round, an unexecuted limit order is deactivated with a probability of five percent. In the second trading round, demand and supply are as follows: 95% of orders from round 1, which where not executed, remain in the market, and new demand and supply arrives. Suppose that the market platform charges hidden transaction costs.

In the first trading round, there is no built-up demand and supply, and thus baseline spread is optimal, that is, $\sigma^1 = \sigma^0$. In the second and third round, the platform profits from built-up excess demand by successively tightening the spread, that is, $\sigma^3 < \sigma^2 < \sigma^1 = \sigma^0$. In the fourth trading round, it becomes optimal to clear at the baseline spread again. Thus, in line with Theorem 3.2, a first price cycle of length 3 evolves already in the first four trading rounds, given by $C^1 = (\sigma^1, \sigma^2, \sigma^3)$. Figure 2 shows the evolution of demand and supply, as well as the optimal bid-ask spread for the first four trading rounds.



((a)) 1^{st} **round:** $P_b^1 = 1.75$, $P_s^1 = 1.25$, $\sigma^1 = 0.5$ ((b)) 2^{nd} **round:** $P_b^2 \approx 1.69$, $P_s^2 \approx 1.31$, $\sigma^2 \approx 0.38$



((c)) 3^{rd} **round:** $P_b^3 \approx 1.65$, $P_s^2 \approx 1.35$, $\sigma^2 \approx 0.3$ ((d)) 4^{th} round: $P_b^4 = 1.75$, $P_s^4 = 1.25$, $\sigma^4 = 0.5$

Figure 2: **Short-run dynamics for hidden TCs.** This figure shows the evolution of demand and supply, buy and sell prices, and the bid-ask spread for the first four trading rounds in a market with linear demand and supply, departure rate $1 - \epsilon = 0.05$.

However, in the second cycle, starting at trading round four, a different pattern emerges. While it indeed holds that $\sigma^5 = \sigma^2$, $\sigma^6 \neq \sigma^3$. Instead, σ^6 is equal to a new global minimum. This is again in line with Theorem 3.1. Each time a new spread is realized, it has to be a new global minimum. Figure 3 shows the long-run evolution of the market.

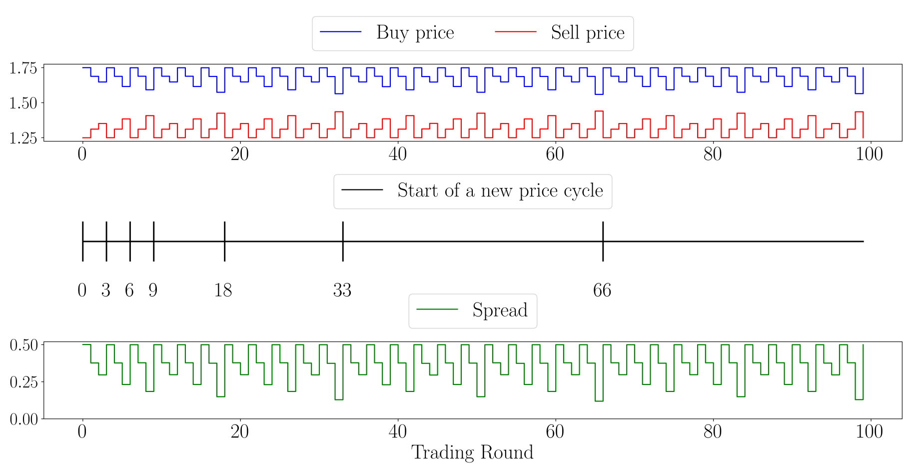


Figure 3: **Long-run dynamics for hidden TCs.** Evolution of buy and sell prices *(Top.)* and spread *(Bottom.)* in a market with linear demand and supply, departure rate $1 - \epsilon = 0.05$. Price cycles emerge with $\limsup_{t \geq 1} \sigma^t = \sigma^0 = 0.5$ and $\liminf_{t \geq 1} \sigma^t \approx 0.12$.

In the first 100 trading rounds, eight different spreads realize, see Example 5.2 below. Suppose they are ordered as $\sigma^0 < \sigma^1 < ... < \sigma^7$. Let $C_i| \cup \sigma$ denote a price cycle C_i, which was broken at a certain point by a new minimum spread σ. In line with Theorem 3.2, we find a recurrent pattern, that is, for each C_n, it starts as a concatenation of $\bigcup_{i=1}^{n-1}$, until it is broken by a new global minimum:

Price Cycle	C_1	C_2	C_3	C_4	C_5	C_6
Recurrent pattern	$\sigma^0 \cup \sigma^1 \cup \sigma^2$	$C_1| \cup \sigma^3$	$C_1| \cup \sigma^4$	$\bigcup_i^2 C_i \cup C_3|\sigma^5$	$\bigcup_i^3 C_i \cup C_4|\sigma^6$	$\bigcup_i^4 C_i \cup C_4|\sigma^7$

4 Transparent TCs lead to Stable Prices

In the following section, I study a market platform charging transparent transaction costs.

Assumption (Ex-Ante Commitment). I assume that the market platform commits to transparent transaction costs ex-ante, that is, it is not allowed to switch to different transparent transaction costs after some round t.

This assumption ensures ex-post individual rationality. Once a trader submits an order, their net value adjustment is based on the existing transparent transaction costs. If the platform were to alter the transaction cost model, the net value adjustment would become inaccurate. Furthermore, this assumption aligns with real-world practices, where platforms typically have to maintain consistent transaction cost structures during clearing windows.

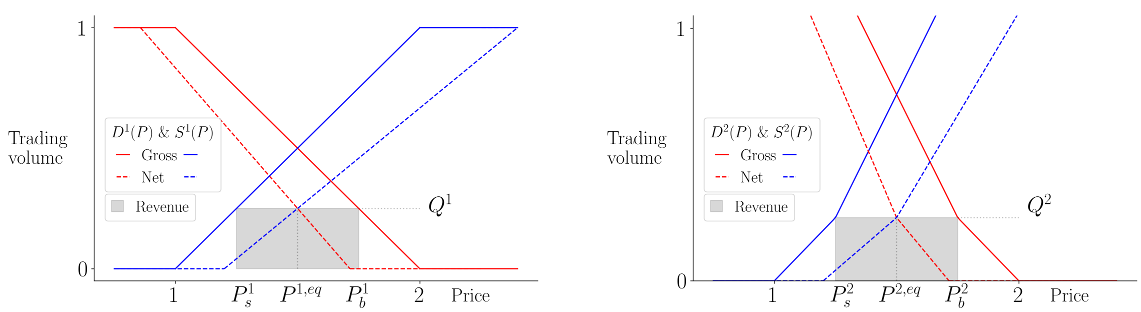
The following theorem shows that the build-up of excess demand and supply, which lead to the predictable cyclical price pattern for hidden transaction costs, does not have the same impact for transparent transaction costs. Instead, for transparent transaction costs, the spread remains constant in the dynamic market. Thus, the predictable cyclical behavior that emerges if the platform charges hidden transaction costs (see Theorem 3.1) cannot be explained by the dynamic market environment on its own.

Theorem 4.1 (Transparent TCs ⇒ Stable Prices)**.** *If a market platform charges transparent transaction costs, the market-clearing price and the bid-ask spread are constant over time.*

Theorem 4.1 has interesting implications. If the market platform commits to transparent transaction costs, the spread does not change over time. Hence, no excess demand and supply is ever cleared, and the revenue is constant over time. This implies that when it comes to revenue-maximization, there is no distinction between myopically maximizing per-round revenue, or maximizing the long-run average. Hence, it is sufficient to set transparent transaction costs to implement the optimal baseline spread σ^0. Jantschgi et al. (2023) show that any transaction cost structure can be linearly scaled, e.g. tuning the percentage of a price fee, to achieve this spread.

Intuition. Why does the build-up of unexecuted orders not influence the market-clearing price? The reason is that transparent transaction costs are added ex-post, and the set of executed orders is calculated solely based on the market-clearing price. This has the following effect: Traders are responsible to adjust to the net value to guarantee ex-post individual rationality. Thus, the market price is not calculated with respect to true demand and supply, but net demand and supply. This corresponds to a market environment without transaction costs, where all profitable orders are cleared. And in such a model, excess demand below the market-clearing price and excess supply above the market-clearing price have no influence on determining the price. Intuitively, net demand and supply become steeper after their intersection, that is, the unique market clearing price. The actual market-clearing price with respect to true values might change, but given the net value considerations, this is never realized in the market. Hence, the market outcome is stationary, even though excess demand and supply builds up. The formal proof is relegated to Appendix B.4.

Example 4.2 (Transparent TCs in a market with linear demand and supply)**.** Suppose that the incoming distribution of demand and supply is linear, see Example 2.2, and that traders with unexecuted orders depart at a rate $1 - \epsilon = 0.05$. That is, before the second trading round, an unexecuted limit order is deactivated with a probability of five percent. Suppose that the market platforms set their transaction costs to maximize revenue, that is, e.g, charges a price fee $p = \frac{1}{6}$. In the first trading round, the revealed net demand are thus given by $D^1(P) = 2 - (1+p)P$ and the revealed net supply is $S^1(P) = (1-p)P - 1$. The baseline market price is $P^{1,eq} = 1.5$. In the second trading round, excess demand has built-up for true values below 1.75 and excess supply has built up for true values above 1.25. However, this corresponds to built-up revealed demand below 1.5 and built-up revealed supply above 1.5. Thus, revealed demand and supply still intersect at $P^{2,eq} = 1.5$, see RHS of Figure 4.



((a)) **First round.** ((b)) **Second round.**

Figure 4: **Market dynamics for a transparent price fee $p = 1/6$.** This figure shows the evolution of demand and supply, buy and sell prices, and the bid-ask spread for the first two trading rounds in a market with linear demand and supply, departure rate $1 - \epsilon = 0.05$, and a transparent price fee $p = 1/6$. The equilibrium price, and thus the spread, remain constant, as excess demand only builds up above the realized market-clearing price and excess supply only builds up below the realized market-clearing price.

This remains true for every trading round $t \geq 1$. For net demand and supply, excess only builds up after their intersection, proving that the market-clearing price remains constant over time. As buy and sell prices are deterministic functions of the market-clearing price, the sequence of realized prices and thus also the sequence of realized spreads remains constant.

5 Platform Competition

In this section, I study competition between a platform charging transparent transaction costs, and a platform charging hidden transaction costs. Traders are faced with the decision, whether to join a platform with transparent transaction costs and associated spread σ^T or join a platform with hidden transaction costs and associated spread σ^H. Which platform should a trader join? The answer to this question depends on the risk attitude of traders and their belief about the market environment.

First, traders derives some utility $u(\sigma)$ from their order being executed at spread σ. A lower spread corresponds to a lower buy price and a higher sell price (Lemma B.2). Thus, a lower spread is more favourable for all traders. In this section, I omit the dependence of the utility function on a trader's true value v_i. While it is true that a smaller spread is more favourable to all traders, certain spreads are not *individually rational* for a trader. Recall that the true value of a trader i is the price, at which they are indifferent between trading and not trading. That is, if the buy price for buyers or the sell price for sellers coincides with their true value v_i, their utility is equal to zero. In the following subsection, I will focus on traders with the most profitable gross values: As we assumed that demand and supply are distributed in the interval $[1, 2]$, a spread of 1 corresponds to a buy price $P_b = 2$ and a sell price $P_s = 1$. In Equation (Utility functions), the utility is equal to zero, if the spread is equal to 1. Hence, these are the utility functions for a buyer with value 2, and a seller with value 1, for whom trade is always individually rational at any spread. In Section 6, I will extend the utility functions for traders with less profitable gross values, for whom trade is not always individually rational.

Assumption (Constant Absolute Risk Aversion). For analytical simplicity, I assume that traders have *constant absolute risk aversion*:

$$u_\alpha(\sigma) = \begin{cases} \frac{1-\exp(-\alpha\cdot(1-\sigma))}{\alpha}, & \text{if } \alpha \neq 0, \\ 1-\sigma, & \text{if } \alpha = 0. \end{cases} \qquad \text{(Utility functions)}$$

This corresponds to *exponential utility functions* with *risk aversion* α, where the term $(1-\sigma)$ corresponds to the fact that smaller spreads yield a higher utility.

This assumption allows for clean comparative statics results with respect to a single parameter α. The proof of Theorem 5.1 can be adjusted to any one-parameter family of utility functions, which monotonically increases or decreases risk-aversion. Moreover, in Appendix A.3, I extend the analysis to markets where traders have different levels of risk aversion $\alpha \in [-M, +M]$.

In order to decide, which market to enter, traders form beliefs about the market environment.

Assumption (Understanding the Market Evolution). Traders know the primitives of the arrival and departure process, and thus understand the market evolution for transparent transaction costs (Theorem 4.1) and for hidden transaction costs (Theorem 3.2). Suppose that the previous bid-ask spreads are not announced publicly.

- *Transparent transaction costs.* Trader i has deterministic beliefs about the spread, that is, they believe that the spread is $\sigma^T \geq 0$. Thus, the utility of joining the platform with transparent transaction costs is equal to $u_\alpha(\sigma^T)$.
- *Hidden transaction costs.* Trader i forms beliefs about how likely spread σ^j is, if the entire sequence of spreads is $(\sigma^j)_{j\geq 1}$. Suppose that beliefs are truncated such that trader i's beliefs have support on a finite set $(\sigma^j)_{1\leq j\leq k}$, and that the departure rate is sufficiently low, such that $k \geq 2$. Let σ^H be a random variable according to this belief. Thus, the utility of joining the platform with hidden transaction costs is $\mathbb{E}_i[u_\alpha(\sigma^H)]$.

The assumption that traders know the spread in the market with transparent transaction costs, is in line with practice. On stock exchanges, traders see the current best available price, and are thus well informed about the current market. The fact that in market with hidden transaction costs, traders are assumed to only form beliefs about the market, but not know the current prices, is also reasonable for some market platforms, as in Payment for Orderflow models, zero-commission platforms sell orders to market makers, who then execute trades at their own discretion. Thus, traders do not the terms of their trade in advance, other than it must meet their limit order. Moreover, in Section 6, I will introduce a rational model of how forward-looking traders form this expecation, based on the idea of *random arrival times*.

Assumption (Utility Maximization). Suppose that trader i aims to maximize their expected utility. Thus, trader i prefers a platform charging transparent transaction costs and spread σ^T over a platform charging hidden transaction costs and spread σ^H, if $u_i(\sigma^T) > \mathbb{E}_i[u_i(\sigma^H)]$. If $u_i(\sigma^T) < \mathbb{E}_i[u_i(\sigma^H)]$, trader i prefers the market charging hidden transaction costs.

Evaluating the market with hidden transaction costs according to $\mathbb{E}_i[u_i(\sigma^H)]$ either corresponds to a social planner, evaluating the overall markets, or to a trader that evaluates the current market conditions, that is, a trader acting as a price taker. This behavior maximizes the expected utility, if traders are perfectly impatient. In Section 6, I extend the analysis and consider patient traders, who do not act as price-takers.

I say that the two markets are in *equilibrium*, if

$$u_\alpha(\sigma^T) = \mathbb{E}_i[u_\alpha(\sigma^H)]. \tag{Market Equilibrium}$$

That is, the trader is indifferent between the two platforms. The following theorem examines the effects of platform competition by analyzing equilibrium properties.

Theorem 5.1 (Platform Competition)**.** *For any risk attitude, there exists a unique equilibrium between a market platform charging transparent transaction costs and a market platform charging hidden transaction costs. Moreover, for a sufficiently low departure rate, in equilibrium:*

- ***Sub-Monopoly Pricing.*** *Transparent transaction costs are charged below the optimal monopolist level, that is,* $0 < \sigma^T < \sigma^0$*.*
- ***Comparative Statics.*** *The spread* σ^T *associated with transparent transaction costs is strictly increasing in the risk-aversion* α*, while the long-run average efficiency is decreasing in* α
- ***Efficiency Crossover.*** *At low levels of risk aversion, transparent transaction costs lead to higher long-run average efficiency, whereas at high levels of risk aversion, hidden transaction costs lead to a higher long-run average efficiency.*

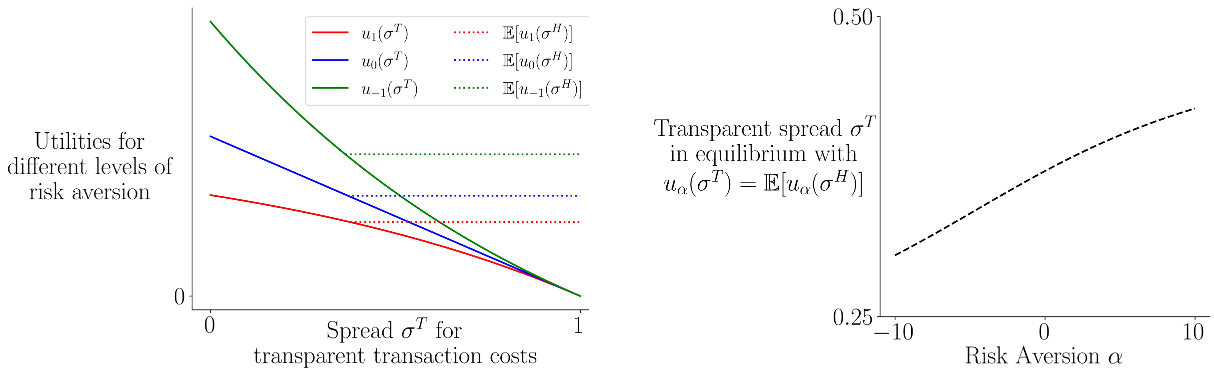
Intuition. A more volatile market with hidden transaction costs offers additional profit opportunities without added risk. Assuming transparent transaction costs lead to the optimal spread σ^0, in a baseline market with no excess demand or supply, the spreads for both transparent and hidden costs coincide (see Proposition 2.1). If the departure rate is sufficiently low, hidden costs will eventually result in spreads $\sigma^t < \sigma^0$. However, Theorem 3.1 indicates that hidden costs never exceed σ^0. Therefore, the spread with hidden costs is always less than or equal to the baseline spread σ^0, and with some probability, it is strictly less. Consequently, traders, regardless of risk aversion, prefer the market with hidden costs. Comparative statics are straightforward: More risk-averse traders prefer the security of transparent costs and a stable spread. Thus, in equilibrium, transparent costs can be set higher for these traders, leading to reduced market efficiency. As risk aversion increases, the transparent spread converges to σ^0, while for risk-loving traders, it converges to the global minimum spread of hidden costs. On average, the efficiency of the market with hidden costs lies between these extremes. Proof details are relegated to Appendix B.5.

Example 5.2 (A trader's choice in a market with linear demand and supply)**.** Consider again a market with linear demand and supply, that is and departure rate $1 - \epsilon = 0.05$. Trader i has exponential utility with constant risk aversion $\alpha \in [-M, M]]$, that is, $u_i(\sigma) = \frac{1-\exp -\alpha\cdot(2-\sigma)}{\alpha}$. Joining the market platform with transparent transaction costs and spread σ^T leads to utility $u_\alpha(\sigma^T)$. Suppose that trader i forms beliefs according to the distribution of spreads in the first 100 trading rounds, see Example 3.3.

	σ^0	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7
Realized spread σ^j	0.5	0.38	0.3	0.225	0.185	0.15	0.13	0.12
$\mathbb{P}[\sigma^H = \sigma^j]$	34%	32%	14%	9%	5%	3%	2%	1%

Table 1: **Trader i's beliefs about price-cycles.**

The utility of joining the market platform with hidden transaction costs and spread σ^H is thus given by $\mathbb{E}[u_\alpha(\sigma^H)] = \sum_{j=0}^{8} u_\alpha(\sigma^j)\mathbb{P}[\sigma^H = \sigma^j]$. Figure 5 shows the comparison of the utilities for different values of the transparent spread σ^T and computes the equilibrium value as a function of the trader's risk attitude. In line with Theorem 5.1, the equilibrium value of the transparent spread is strictly increasing in the risk attitude α.



((a)) **Risk-neutral trader.** ((b)) **Risk-averse trader**.

Figure 5: **Equilibrium Analysis**. This figure compares the expected utility of a trader for joining either a platform with transparent transaction costs and spread σ^T or a platform with hidden transaction costs and spread σ^H in the market with linear demand and supply. **Left.** Comparison of utilities for different choices of $\sigma^T \in [0,1]$ for a risk-averse trader with $\alpha = 0$ (red), a risk-neutral trader with $\alpha = 0$ (blue), and a risk-loving trader with $\alpha = -1$ (green). The intersection gives the equilibrium value of the transparent spread σ^T. **Right.** The equilibrium value of the transparent spread as a function of a trader's risk attitude.

6 Strategic Market Entry

Up to this point, I have considered traders who are *myopic* — those who act as price-takers, reporting their true valuations upon market entry. This behavior is optimal in a static large market, but also when traders are *perfectly impatient*, with a discount factor of zero, meaning they seek to maximize expected utility in a single trading round. In this section, I shift focus to *patient* traders — those who understand the market's evolution and acknowledge that their orders may not be executed immediately.

Recall from Section 5 that traders have constant absolute risk aversion with a utility function $u_\alpha(\sigma)$, which is strictly decreasing in the spread. I now introduce the following addition to account for their true value: A trader's utility also depends on their true value v_i, writ $u_{v_i}(\sigma)$, where v_i determines the unique point, where a trader is indifferent between trading and not trading. Let $P_b(\sigma)$ and $P_s(\sigma)$ be the buy and sell price associated with a given spread. The utility functions are thus of the following form:

$$u_{v_b}(\sigma) = \begin{cases} \frac{1-\exp(-\alpha(v_b - P_b))}{\alpha}, & \text{if } v_b \geq P_b \wedge \alpha \neq 0, \\ v_b - P_b, & \text{if } v_b \geq P_b \wedge \alpha = 0, \end{cases} \quad u_{v_s}(\sigma) = \begin{cases} \frac{1-\exp(-\alpha(P_s - v_s))}{\alpha}, & \text{if } P_s \geq v_s \wedge \alpha \neq 0, \\ P_s - v_s, & \text{if } P_s \geq v_s \wedge \alpha = 0. \end{cases}$$

If $v_b < P_b(\sigma)$ and $v_s > P_s(\sigma)$, I assume that traders have a negative utility $-L$.[15]

I assume that, upon arrival, traders do not know the current point in the cycle, but model their arrival at the market at a random time.

Assumption (Random Arrival Times). Suppose again that trader i has some beliefs about the evolution of the bid-ask spread $\Sigma = (\sigma^t)_{t\geq 1}$ in the market. Let σ^{inf} denote $\inf_{t\geq 0} \sigma^t$ and σ^{sup} denote $\sup_{t\geq 0} \sigma^t$. Let t_i denote a random variable on $\mathbb{N}_{\geq 1}$. I call this the *random arrival time* of trader i. That is, upon arrival, trader i believes that the current spread is σ^{t_i}.

As argued in Section 5, this assumption is irrelevant for transparent transaction costs, as the spread is constant over time, and thus mirrors the fact that traders know the transparent spread before arriving to the market. For hidden transaction costs, this mirrors the fact that prices are not publicly announced in advance, but sophisticated traders at least know the structure of the price cycles. A second assumption is that traders discount their utility, depending on when their order is executed.

Assumption (Discounted Expected Utility). Suppose that trader i with true value v_i submits a value v_i'. Let $T_i(v_i', \Sigma, t_i)$ denote the first time after a trader's random arrival t_i, where order v_i' is executed at spread $\sigma^{T_i(v_i', \Sigma, t_i)}$:

$$T_b(v_b', \Sigma, t_b) = \inf_{t\geq t_b} \{v_b' \geq P_b^t\} \text{ and } T_s(v_s', \Sigma, t_s) = \inf_{t\geq t_s} \{v_s' \leq P_s^t\} \qquad \text{(Order Execution Time)}$$

The *expected utility* of trader i with *discount factor* $\delta \in (0, 1]$ is

$$\mathbb{E}_{t_i}[\delta^{T_i(v_i', \Sigma, t_i)} \cdot u_{v_i}(\sigma^{T_i(v_i', \Sigma, t_i)})] \qquad \text{(Expected Utility)}$$

Suppose that upon arrival at the platform, traders submit a value to maximize their expected discounted utility. A *best response* $BR(v_i)$ for a trader with true value v_i satisfies

$$BR(v_i) \in \max_{v_i' \in [1,2]} \mathbb{E}_{t_i}[\delta^{T_i(v_i', \Sigma, t_i)} \cdot u_{v_i}(\sigma^{T_i(v_i', \Sigma, t_i)})]. \qquad \text{(Best Response)}$$

I use the convention that reporting a value v_i' that is never executed, results in zero utility, as $T_i(v_i', \Sigma, t_i) = \infty$ in that case.

[15]This assumption is not crucial, and can be extended to any strictly negative function.

6.1 Transparent TCs are Robust to Strategic Market Entry

For transparent transaction costs, patience and sophistication do not alter the market dynamics. Since the spread $\Sigma = (\sigma^t)_{t\geq 1}$ remains stable over time (Theorem 4.1), there is no incentive for individual traders to gain by misreporting their values, as truthfulness is a best response. The market is thus robust against strategic market entry.

Theorem 6.1 (Transparent TCs ⇒ robust against strategic market entry)**.** *For a market platform with transparent transaction costs, truthfulness is a best response.*

Proof. This follows from Theorem 4.1, as the spread $\Sigma = (\sigma^t)_{t\geq 1}$ is constant over time. Hence, truthfully reporting their true value maximizes the expected utility. Either, a trader's order is executed immediately, or never. If the true value leads to no execution, this is also optimal, as reporting an order that would get executed would lead to a negative utility. □

6.2 Hidden TCs are Fragile to Strategic Market Entry

Because hidden transaction costs lead to price cycles (Theorems 3.1 and 3.2), reporting truthfully might not be optimal. Best responses balance the trade-off between maximizing utility by submitting a more aggressive order, which only gets executed at a small and thus more profitable spread versus faster execution by being less aggressive. First, note that it is sufficient to consider reported values that correspond to one of the realized spreads in the sequence of cycles $\Sigma = (\sigma^t)_{t\geq 1}$. For a submitted order v_i', let σ^* be the spread with the closest buy or sell price in Σ, at which v_i' could get executed. Submitting exactly this buy or sell price results in the same expected utility. As this space is countable, it is easy to show that a best response exists.

In the remainder of this section, I will specifically consider *perfectly patient traders* with discount rate δ sufficiently close to 1. It is straightforward to see that for perfectly patient trader and the market evolution $\Sigma = (\sigma^t)_{t\geq 1}$, if σ^{inf} denotes $\inf_{t\geq 0}\sigma^t$, then the best response of buyers and sellers are close to the associated buy and sell prices P_b^{inf} and P_s^{inf}, if this is individually rational given their true value. This is the case, because for δ sufficiently close to 1, the increase in utility from receiving a better buy and sell price outweighs the exponential discounting of receiving that spread at a later point in time. Hence, traders submit an order that only gets executed, when the spread is close to minimal. If this limit is not individually rational, suppose that traders report truthfully, because then, no individually rational order will ever be executed. I call this strategy the *patient best response*:

$$PBR(v_b,\Sigma)=\begin{cases}P_b^{inf} & \text{if } v_b \geq P_b^{inf}\\ v_b & \text{if } v_b < P_b^{inf}\end{cases} \quad PBR(v_s,\Sigma)=\begin{cases}P_s^{inf} & \text{if } v_s \leq P_s^{inf}\\ v_s & \text{if } v_s > P_s^{inf}\end{cases} \qquad \text{(Patient Best Responses)}$$

If all traders use patient best responses, revealed demand and supply adjust as follows:

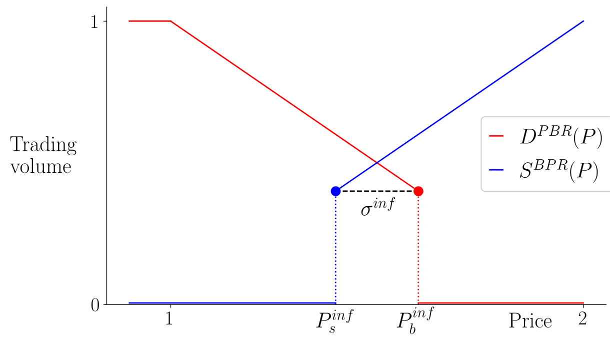


Figure 6: **Demand and Supply for Patient Best Responses.** This figure shows the patient best response adjustments to linear demand and supply. Up until the buy and sell price associated to σ^{inf}, demand and supply are zero, as there are no buy and sell orders with such a value. At these buy prices, demand and supply jump up to the true demand and supply curve.

Level k Reasoning. Research suggests that humans often engage in iterative reasoning to anticipate the actions of others, with the depth of this reasoning varying depending on the individual and the context, see Crawford (2019) for an overview. In the following, I apply the concept of *level k reasoning* to the patient best responses of traders.

Suppose that, as in Section 3, all traders truthfully report their values. Let be $\Sigma_0 = (\sigma_0^t)_{t\geq 1}$ the evolution of the bid-ask spread according to Theorem 3.1 and Theorem 3.2. I call Σ^0 the *level 0 market evolution*. Denote the patient best response trading strategies against Σ_0 by PBR_1, that is, reporting the buy and sell prices associated with σ_0^{inf}, if it is individually rational, and being truthful otherwise. I call these trading strategies *level 1 best responses*, as they best respond to the market evolution for unsophisticated traders. Let D^{PBR_1} and S^{PBR_1} be the adjustments of demand and supply, if traders submit according to PBR_1.

Suppose that $D^{PBR_1}(\cdot)$ and $S^{PBR_1}(\cdot)$ are the new incoming value distributions in each trading round. It follows analogous to Theorem 3.1 and Theorem 3.2 that a market platform charging hidden transaction costs would create price cycles according to a sequence of spreads Σ_{PBR_1}. I call Σ_{PBR_1} the *level 1 market evolution* and denote by σ_1^{inf} the new infimal spread. Note that the new baseline spread in Σ_{PBR_1} must be smaller or equal to σ_0^{inf}, as any spread greater than σ_0^{inf} would lead to zero revenue, see Figure 6. Forward looking and patient traders could know about the price cycles according to Σ_{PBR_1}. Then their best response becomes to bid the new infimal spread $\sigma_1^{inf} < \sigma_0^{inf}$. Let PBR_2 be the patient best responses to Σ_{PBR_1}. I call these trading strategies *level 2 best responses*, as they best respond to the market evolution, where all traders use level 1 best responses. Let D^{BR_2} and S^{BR_2} be the associated adjustments to demand and supply for level 2 best responses. Let Σ_{BR_2} be the evolution of spreads and associated price cycles with respect to D^{BR_2} and S^{BR_2}, called the *level 2 market evolution.* Applying this reasoning iteratively yields *level k best responses BR_k* as best responses to the *level $k-1$ market evolution* $\Sigma_{BR_{k-1}}$ with $\sigma_k^{inf} > \sigma_{k-1}^{inf}$.

Next, I will show that strategic market entry in the form of level k-best responses can lead to a form of market failure:

Assumption (Minimal Revenue Requirement). Suppose that a market platform needs a fixed amount of revenue $R > 0$ per round to remain profitable, which is satisfied for baseline demand and supply, that is $R(\sigma^0) \geq R$. I call R the *profitability threshold* and say that *market failure* occurs, if in some trading round, the profit is less or equal to R.

Note that, if clearing the market without excess is profitable, then, even if price-cycles occur, the market remains profitable. That is, because price-cycles always increase revenue. Hence the baseline spread, where no excess is cleared, is a lower bound on revenue.

For many market platforms, the minimal revenue assumption is reasonable because they typically incur significant overhead costs associated with operating the exchange, such as infrastructure, regulatory compliance, and transaction processing. If the spread and trading volume are too small, the revenue generated may not cover these fixed costs, making it unprofitable for the platform to continue operating. In such cases, it may not be economically viable to execute trades, as doing so would fail to meet the minimum revenue needed to sustain the platform's operations.

As soon as price cycles occur, strategic market entry can lead to market failure, even for level 2 best responses, if the profitability threshold R is close to the benchmark revenue for baseline demand and supply, that is, if the platform is not generating a lot of revenue to begin with. However, a stronger statement holds: No matter the profitability threshold R, market failure occurs for sufficiently patient traders.

Theorem 6.2 (Hidden TCs ⇒ fragile against strategic market entry)**.** *Consider a market platform charging hidden transactions. For any profitability threshold $R > 0$, if traders are sufficiently patient and have a sufficiently low departure rate, there exists $k \geq 1$, such that market failure occurs in the level k-market evolution.*

Intuition. If traders adopt level k-best responses, the baseline spread in the level k market, $\sigma^{\sup,k}$, will strictly decrease over time because $\sigma^{\inf,k} < \sigma^{\sup,k}$ and $\sigma^{\sup,k+1} \leq \sigma^{\inf,k}$. Therefore, it suffices to show that, eventually, there will exist a k such that the baseline spread in the level k market evlution becomes unprofitable. If the departure rate is sufficiently low, price cycles will be significant, causing the smallest spread in the cycle to become very small. If this spread becomes so small that clearing the baseline market without excess at this spread is no longer profitable, then market failure occurs in the level $k+1$ market, as the new baseline spread will be at this level or even smaller. For every profitability threshold $R > 0$, let σ^* be the smallest spread such that clearing a market without excess remains profitable. If the departure rate is sufficiently low, then at some point, a smaller spread will be realized to clear the excess, leading to market failure in the next period. Proof details are provided in Appendix B.6.

Thus, Theorem 6.2 reveals a close connection to the Coase conjecture (Coase 1960), where a platform's failure to commit to stable transaction costs leads to market failure through strategic market entry of traders. However, the crucial difference to classical Coasian dynamics is that, as prices are not revealed publicly, traders have to time their purchase through strategically submitting aggressive limit orders.

Hence, while transparent transaction costs are robust to transparent transaction costs, market platforms with hidden transaction costs need unsophisticated traders, as the market could significantly fail otherwise.

7 Conclusion

This paper studied the effects of transaction cost transparency on market dynamics. In a market with homogeneous arrival and memoryless departure, I have shown that transparency is a crucial necessity for market stability and efficiency. While transparent transaction costs lead to a stable market outcome, hidden transaction costs lead to additional volatility and predictable price cycles . In a model of platform competition, I showed that platforms with transparent transaction costs must lower them to compete with the profit opportunities of more volatile markets with hidden costs. An efficiency crossover emerges: at higher risk aversion, transparent costs are less efficient on average, while the reverse holds for lower risk aversion. This creates a natural market segmentation, with less risk-averse traders favoring hidden costs for higher returns, and more risk-averse traders preferring the stability of transparent costs. Additionally, I uncovered a critical vulnerability in platforms with hidden transaction costs when faced with strategic market entry. While transparent transaction costs are robust against such strategies, hidden costs are fragile due to a newly identified Coasian dynamic that can lead to market failure. This fragility underscores the necessity for platforms with hidden transaction costs to attract unsophisticated traders to maintain their long-term viability.

Given that hidden transaction costs are integral to the operations of online trading platforms and market makers, these findings align with regulatory concerns, emphasizing the need for caution in adopting zero-commission trading models without fully understanding their broader market implications.

There are several promising avenues for future research. One key question is whether empirical evidence can be found—beyond the well-established link to wider bid-ask spreads—that connects zero-commission models to increased market volatility and cyclical pricing. Additionally, given that both transparent and hidden transaction costs are charged in practice, what explains the choice in a particular market in practice? Another intriguing question is whether hidden transaction costs, which create profit opportunities through strategic market entry, could be systematically exploited by trading algorithms or sophisticated human traders. This raises the question of how access to historical pricing data and real-time order book information might influence these strategies, emphasizing the critical relationship between transaction cost transparency and order book visibility. For platforms operating under a zero-commission model, optimizing or controlling the flow of information could become a strategic necessity, further reinforcing concerns about intransparency. Lastly, how might these findings inform regulatory bodies and policymakers in addressing the potential risks posed by hidden transaction costs and the strategic behaviors they incentivize?

References

Anand, Amber, Jian Hua, and Tim McCormick (2016), "Make-take structure and market quality: Evidence from the us options markets." *Management Science*, 62, 3271–3290.

Armstrong, Mark (2006), "Competition in two-sided markets." *RAND Journal of Economics*, 37, 668–691.

Aumann, Robert (1964), "Markets with a continuum of traders." *Econometrica*, 32, 39–50.

Banerjee, Abhijit V (1992), "A simple model of herd behavior." *The Quarterly Journal of Economics*, 107, 797–817.

Barclay, Michael J, Eugene Kandel, and Leslie M Marx (1998), "The effects of transaction costs on stock prices and trading volume." *Journal of Financial Intermediation*, 7, 130–150.

Battalio, Robert and Craig W Holden (2001), "A simple model of payment for order flow, internalization, and total trading cost." *Journal of Financial Markets*, 4, 33–71.

Battalio, Robert, Andriy Shkilko, and Robert Van Ness (2016), "To pay or be paid? the impact of taker fees and order flow inducements on trading costs in us options markets." *Journal of Financial and Quantitative Analysis*, 51, 1637–1662.

Battalio, Robert H and Tim Loughran (2008), "Does payment for order flow to your broker help or hurt you?" *Journal of Business Ethics*, 80, 37–44.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), "A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of Political Economy*, 100, 992–1026.

Board, Simon and Marek Pycia (2014), "Outside options and the failure of the coase conjecture." *American Economic Review*, 104, 656–671.

Board, Simon and Andrzej Skrzypacz (2016), "Revenue management with forward-looking buyers." *Journal of Political Economy*, 124, 1046–1087.

Brzustowski, Thomas, Alkis Georgiadis-Harris, and Balázs Szentes (2023), "Smart contracts and the coase conjecture." *American Economic Review*, 113, 1334–1359.

Budish, Eric, Peter Cramton, and John Shim (2015), "The high-frequency trading arms race: Frequent batch auctions as a market design response." *The Quarterly Journal of Economics*, 130, 1547–1621.

Bulow, Jeremy and Paul Klemperer (1994), "Rational frenzies and crashes." *Journal of Political Economy*, 102, 1–23.

Cantillon, Estelle and Pai-Ling Yin (2011), "Competition between exchanges: A research agenda." *International journal of industrial organization*, 29, 329–336.

Chen, Daniel and Anthony Lee Zhang (2020), "Subsidy schemes in double auctions." *Available at SSRN 3668394*.

Chordia, Tarun and Avanidhar Subrahmanyam (1995), "Market making, the tick size, and payment-for-order flow: theory and evidence." *Journal of Business*, 543–575.

Coase, Ronald (1960), "The problem of social cost." *Journal of Law and Economics*, 3, 1–44.

Coase, Ronald H (1972), "Durability and monopoly." *The Journal of Law and Economics*, 15, 143–149.

Colliard, Jean-Edouard and Thierry Foucault (2012), "Trading fees and efficiency in limit order markets." *Review of Financial Studies*, 25, 3389–3421.

Conlisk, John, Eitan Gerstner, and Joel Sobel (1984), "Cyclic pricing by a durable goods monopolist." *The Quarterly Journal of Economics*, 99, 489–505.

Crawford, Vincent P (2019), "Experiments on cognition, communication, coordination, and cooperation in relationships." *Annual Review of Economics*, 11, 167–191.

Demsetz, Harold (1968), "The cost of transacting." *Quarterly Journal of Economics*, 82, 33–53.

Eckert, Andrew (2002), "Retail price cycles and response asymmetry." *Canadian Journal of Economics/Revue canadienne d'économique*, 35, 52–77.

Ernst, Thomas and Chester S Spatt (2022), "Payment for order flow and asset choice." Technical report, National Bureau of Economic Research.

European Securities and Markets Authority (2021), "Esma public statement on payment for order flow (pfof) and zero-commission brokers." URL https://www.esma.europa.eu/sites/default/files/library/esma35-43-2749_esma_public_statement_pfof_and_zero-commission_brokers.pdf.

Expert Market Research (2023), "Global online trading platform market share, size, growth, analysis, trends, forecast: 2024-2032." Technical report. Accessed: insert date here.

Financial Conduct Authority (2012), "Guidance on the practice of 'payment for order flow'." URL https://www.fca.org.uk/publication/finalised-guidance/fg12-13.pdf.

Foucault, Thierry, Ohad Kadan, and Eugene Kandel (2013), "Liquidity cycles and make/take fees in electronic markets." *Journal of Finance*, 68, 299–341.

Glosten, Lawrence and Paul Milgrom (1985), "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders." *Journal of Financial Economics*, 14, 71–100.

Gul, Faruk, Hugo Sonnenschein, and Robert Wilson (1986), "Foundations of dynamic monopoly and the coase conjecture." *Journal of economic Theory*, 39, 155–190.

Hollifield, Burton, Robert A Miller, Patrik Sandås, and Joshua Slive (2006), "Estimating the gains from trade in limit-order markets." *The Journal of Finance*, 61, 2753–2804.

Hörner, Johannes and Larry Samuelson (2011), "Managing strategic buyers." *Journal of Political Economy*, 119, 379–425.

Jantschgi, Simon, Heinrich Nax, Bary Pradelski, and Marek Pycia (2022), "Double auctions and walrasian equilibrium." *Available at SSRN 4043532*.

Jantschgi, Simon, Heinrich Nax, Bary Pradelski, and Marek Pycia (2023), "Markets and transaction costs." *Available at SSRN 4043576*.

Kandel, Eugene and Leslie M Marx (1999), "Payments for order flow on nasdaq." *The Journal of Finance*, 54, 35–66.

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

Lee, In Ho (1998), "Market crashes and informational avalanches." *The Review of Economic Studies*, 65, 741–759.

Loertscher, Simon, Ellen V Muir, and Peter G Taylor (2022), "Optimal market thickness." *Journal of Economic Theory*, 200, 105383.

Madhavan, Ananth (2000), "Market microstructure: A survey." *Journal of financial markets*, 3, 205–258.

Malinova, Katya and Andreas Park (2015), "Subsidizing liquidity: The impact of make/take fees on market quality." *Journal of Finance*, 70, 509–536.

Maskin, Eric and Jean Tirole (1988), "A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles." *Econometrica: Journal of the Econometric Society*,

571–599.
McAfee, R Preston and Thomas Wiseman (2008), "Capacity choice counters the coase conjecture." *The Review of Economic Studies*, 75, 317–331.
Mendelson, Haim (1982), "Market behavior in a clearing house." *Econometrica: Journal of the Econometric Society*, 1505–1524.
Nisan, Noam (2023), "Serial monopoly on blockchains."
Noel, Michael D (2007), "Edgeworth price cycles, cost-based pricing, and sticky pricing in retail gasoline markets." *The Review of Economics and Statistics*, 89, 324–334.
Noel, Michael D (2008), "Edgeworth price cycles and focal prices: Computational dynamic markov equilibria." *Journal of Economics & Management Strategy*, 17, 345–377.
Noussair, Charles, Stephane Robin, and Bernard Ruffieux (1998), "The effect of transaction costs on double auction markets." *Journal of Economic Behavior & Organization*, 36, 221–233.
O'Hara, Maureen (1998), *Market microstructure theory*. John Wiley & Sons.
O'Hara, Maureen and George S Oldfield (1986), "The microeconomics of market making." *Journal of Financial and Quantitative analysis*, 21, 361–376.
Parlour, Christine A and Uday Rajan (2003), "Payment for order flow." *Journal of Financial Economics*, 68, 379–411.
Rochet, Jean-Charles and Jean Tirole (2003), "Platform competition in two-sided markets." *Journal of the European Economic Association*, 1, 990–1029.
Securities and Exchange Commission (2022), "Sec regulation best execution." URL https://www.sec.gov/files/rules/proposed/2022/34-96496.pdf.
Shiller, Robert (1995), "Conversation, information, and herd behavior." *The American Economic Review*, 85, 181–185.
Sobel, Joel (1991), "Durable goods monopoly with entry of new consumers." *Econometrica: Journal of the Econometric Society*, 1455–1485.
Stokey, Nancy L (1981), "Rational expectations and durable goods pricing." *The Bell Journal of Economics*, 112–128.
Tan, Guofu and Junjie Zhou (2021), "The effects of competition and entry in multi-sided markets." *The Review of Economic Studies*, 88, 1002–1030.
Tatur, Tymon (2005), "On the trade off between deficit and inefficiency and the double auction with a fixed transaction fee." *Econometrica*, 73, 517–570.
Teh, Tat-How, Chunchun Liu, Julian Wright, and Junjie Zhou (2023), "Multihoming and oligopolistic platform competition." *American Economic Journal: Microeconomics*, 15, 68–113.
U.S. Securities and Exchange Commission (2005), "Regulation national market system." Available at https://www.sec.gov/rules/final/34-51808.pdf.
van Ryzin, Garrett J and Kalyan T Talluri (2005), "An introduction to revenue management." In *Emerging theory, methods, and applications*, 142–194, Informs.
Wang, Zhongmin (2009), "(mixed) strategy in oligopoly pricing: Evidence from gasoline price cycles before and under a timing regulation." *Journal of Political Economy*, 117, 987–1030.
Weyl, E Glen (2010), "A price theory of multi-sided platforms." *American Economic Review*, 100, 1642–1672.

A Additional Results

A.1 Evolution of Demand and Supply

Lemma A.1 (Evolution of Demand and Supply)**.** *Consider a market platform charging hidden transaction costs. Fix $P_b < P_b^{'}$ and $P_s > P_s^{'}$.*

1. *For all $t > 1$ we have that*

$$\begin{aligned} D^t(P_b) - D^t(P_b^{'}) &\leq \frac{1-\epsilon^{t+1}}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) \leq \frac{1}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) \\ S^t(P_s) - S^t(P_s^{'}) &\leq \frac{1-\epsilon^{t+1}}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right) \leq \frac{1}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right) \end{aligned} \tag{1}$$

2. *For all T and $t > T$*

$$\begin{aligned} D^t(P_b) - D^t(P_b^{'}) &\leq \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) + \epsilon^{t-T}\left(Z_D^T(P_b) - Z_D^T(P_b^{'})\right) \\ S^t(P_s) - S^t(P_s^{'}) &\leq \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right) + \epsilon^{t-T}\left(Z_S^T(P_s) - Z_S^T(P_s^{'})\right). \end{aligned} \tag{2}$$

3. *For all T and $t > T$, if for all t' with $T < t' < t$ we have that $P_b^{t'} \geq P_b^{'} > P_b$, then*

$$\begin{aligned} D^t(P_b) - D^t(P_b^{'}) &= \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) + \epsilon^{t-T}\left(Z_D^T(P_b) - Z_D^T(P_b^{'})\right) \\ S^t(P_s) - S^t(P_s^{'}) &= \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right) + \epsilon^{t-T}\left(Z_S^T(P_s) - Z_S^T(P_s^{'})\right). \end{aligned} \tag{3}$$

4. *For all T such that $P_b^T \leq P_b < P_b^{'}$ (or $T = 0$), and all $t > T$, we have that*

$$\begin{aligned} D^t(P_b) - D^t(P_b^{'}) &\leq \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) \leq \frac{1}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) \\ S^t(P_s) - S^t(P_s^{'}) &\leq \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right) \leq \frac{1}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right). \end{aligned} \tag{4}$$

5. *For all T such that $P_b^T \leq P_b < P_b^{'}$ (or $T = 0$), if for all t' with $T < t' < t$, we also have that $P_b^{t'} \geq P_b^{'} > P_b$, then*

$$\begin{aligned} D^t(P_b) - D^t(P_b^{'}) &= \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) \leq \frac{1}{1-\epsilon}\left(D^0(P_b) - D^0(P_b^{'})\right) \\ S^t(P_s) - S^t(P_s^{'}) &= \frac{1-\epsilon^{t-T}}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right) \leq \frac{1}{1-\epsilon}\left(S^0(P_s) - S^0(P_s^{'})\right). \end{aligned} \tag{5}$$

The proof of Lemma A.1 is relegated to Appendix B.7.

A.2 Hidden Transaction Costs that maximize long-run average revenue

In this subsection, I consider a market platform charging hidden transaction costs, which does not myopically maximizes per-round revenue, but instead chooses an infinite sequence of spreads to maximize the long-run average revenue.

The main result of this Appendix is an extension of Theorem 3.1, demonstrating that the optimal sequence of spreads is oscillatory when the objective is to maximize long-run average revenue. I use the term *oscillatory* rather than *cyclical* because the primary proof presented in this Appendix relies on abstract arguments to show that the sequence of spreads cannot eventually become monotone and must instead fluctuate indefinitely. However, the method used to prove Theorem 3.2, which addresses the predictability of these fluctuations, does not extend to this context, which is why I refer to the sequence as oscillatory rather than cyclical.

To formalize this result, I introduce additional notation: Consider an infinite sequence of spreads $\sigma = (\sigma_{tt \geq 1})$. A sequence of spreads is *oscillatory*, if for any $t \in \mathbb{N}$, there exist $t', t'' > t$, such that $\sigma^{t'} < \sigma^{t'+1}$ and $\sigma^{t''} > \sigma^{t''+1}$. A sequence of spreads is *eventually monotone*, if there exists $t \in \mathbb{N}$, such that (i) for all $t' > t$ it holds that $\sigma^{t'} \leq \sigma^{t'+1}$ or (ii) for all $t' > t$ it holds that $\sigma^{t'} \geq \sigma^{t'+1}$. Note that a sequence of spreads is not oscillatory if and only if it is eventually monotone.

The following theorem shows that any optimal sequence of spreads must be oscillatory.

Theorem A.2 (Hidden TCs + long-run revenue maximization $\Rightarrow$ Price oscillations)**.** *If the departure rate is sufficiently low, any eventually monotone sequence of spreads does not maximize long-run average revenue.*

The proof is relegated to Appendix B.8.

A.3 Platform Competition for Different Risk-Aversions

In the following subsection, I will consider a continuum of traders with different risk attitudes and derive an equilibrium condition for transparent transaction costs to compete against hidden transaction costs.

A market platform with transparent transaction cost must lower the spread to attract traders with lower risk-aversion. On the other hand, for a lower spread, the market platform decreases their revenue per trade. What is the optimal choice for transparent transaction costs to compete against hidden transaction costs?

Distribution of risk attitudes. I assume that traders share the same belief about the market environment, but have different risk attitudes. For analytical simplicity, I again assume that traders have *constant relative risk aversion*. That is, for every trader i, their utility functions is $u_i(\sigma) = \tilde{u}_i(2-\sigma)$, where $\tilde{u}_i(\cdot)$ comes from the family $\mathcal{U}^\alpha = (\tilde{u}_\alpha)_{\alpha\in[-M,+M]}$ of *exponential utility functions* with *risk-aversion* α. That is, $\tilde{u}_\alpha(x) = \frac{1-\exp(-\alpha\cdot x)}{\alpha}$ for $\alpha \neq 1$ and $\tilde{u}_0(x) = x$. Again, a higher α correspond to more risk-averse traders.

I assume that a trader's risk attitude is independent of their value or preferred quantity of trade, and whether they are a buyer or seller: That is, there exists a continuous and strictly positive function $f(\alpha)$ on $[-M, M]$, such that value distributions of buyers and sellers with relative risk aversion α are $f(\alpha)\mu_B^t(\cdot)$ and $f(\alpha)\mu_S^t(\cdot)$. Thus, if buyers and sellers with risk attitude in $\mathcal{A} \subset [-M, M]$ choose to enter a platform, demand $\lambda(\mathcal{A})D^0(\cdot)$ and the supply $\lambda(\mathcal{A})S^0(\cdot)$ arrives, where $\lambda_\mathcal{A} = \int_\mathcal{A} f(\alpha)d\alpha$ denotes the mass of traders with risk attitude in $\mathcal{A}$. The pair $(\mathcal{U}^\alpha, f^\alpha)$ is called the *distribution of risk attitudes*.

Attracting Traders. For a given risk attitude α, a trader will prefer the platform with transparent transaction costs and spread σ^T or the platform with hidden transaction costs and spread σ^H. Let $\mathcal{A}^T(\sigma^T, \sigma^H) = \{\alpha \in [-M, M] : u_\alpha(\sigma^T) \geq \mathbb{E}[u_\alpha(\sigma^H)]\}$ denote the set of risk attitudes, such that a trader with that level of risk-aversion prefers the platform charging transparent transaction costs with spread σ^T. Similarly, $\mathcal{A}^H(\sigma^T, \sigma^H) = \{\alpha \in [-M, M] : u_\alpha(\sigma^T) \leq \mathbb{E}[u_\alpha(\sigma^H)]$ corresponds to traders, who prefer the platform charging hidden transaction costs with spread σ^H.

Market equilibrium. For a platform charging transparent transaction costs with associated spread σ^T, the revenue when facing competition from a platform with hidden transaction cost and associated spread σ^H is given by $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H)) \cdot R^0(\sigma^T)$, where $R^0(\sigma^T)$ is the revenue from spread σ^T in the baseline model with a unit mass of traders on each market side. This holds, because the spread σ^T attracts a total mass of $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H))$ traders on both market sides, and by assumption, the risk aversion is independent of the value distribution. Thus, incoming demand and supply has the same shape, but is multiplied by the mass of incoming traders. Finally, I introduce the following notion of a *market equilibrium*, where a market platform with transparent transaction costs sets the spread σ^T to maximize their revenue when competing against a market platform charging hidden transaction costs with spread σ^H, and thus best responds against the competition. I say that the pair (σ^T, σ^H) is an *equilibrium*, if $\sigma^T \in \arg\max_{\sigma \geq 0} \lambda(\mathcal{A}(\sigma, \sigma^H)) \cdot R^0(\sigma)$.

The following theorem shows that there exists an equilibrium between a platform charging transparent transaction costs and a platform charging hidden transaction costs. Moreover, this equilibrium results in a market segmentation, where less risk-averse traders prefer hidden transaction costs, while more risk-averse traders prefer transparent transaction costs.

Theorem A.3 (Equilibrium existence and properties). *For any distribution of risk attitudes $(\mathcal{U}^\alpha, f^\alpha)$, there exists an equilibrium (σ^T, σ^H) between a market platform charging transparent transaction costs and a market platform charging hidden transaction costs. Moreover:*

- ***Sub-monopoly pricing.*** *Transparent transaction costs are charged below the optimal monopolist level, that is, $0 < \sigma^T < \sigma^0$.*
- ***Market segmentation.*** *Less risk-averse traders enter the platform charging hidden transaction costs, and more risk-averse traders enter the platform charging transparent transaction costs. That is, there exists $\alpha^* \in [-M, M]$ with $\mathcal{A}^H(\sigma^T, \sigma^H) = [-M, \alpha^*]$ and $\mathcal{A}^T(\sigma^T, \sigma^H) = [\alpha^*, M]$.*

Transparent transaction costs lead to a stable market, while hidden transaction costs lead to a volatile market. How can there be an equilibrium?

Intuition. Different traders have different risk attitudes. Suppose that for traders with a fixed risk attitude $\alpha \in [-M, M]$, transparent transaction costs with spread σ^T and hidden transaction costs with spread σ^H are such that $u_\alpha(\sigma^T) = \mathbb{E}[u_\alpha(\sigma^H)]$. Such traders are indifferent between the two platforms. More risk-averse traders strictly prefer the certain spread σ^T over the uncertain spread σ^H. Less risk-averse traders strictly prefer the platform with hidden transaction costs. Thus, for any transparent transaction costs with spread σ^T, the market is segmentating, that is, there exists a threshold risk attitude $\alpha^*(\sigma^T)$, such that traders with lower risk-aversion choose the platform with hidden transaction costs, hoping for a market entry at a currently low spread, while trader with higher risk-aversion choose the platform with transparent transaction costs to avoid market entry at a high spread. Given a distribution of risk attitudes, transparent transaction costs can be tuned to balance the trade-off between trying to attract more traders by offering a lower spread versus maximizing the revenue from a higher spread. More formally, I show that the map $\sigma^T \mapsto \lambda(\mathcal{A}(\sigma^T, \sigma^H)) \cdot R^0(\sigma^T)$ is continuous and thus attains a maximum by the Extreme Value Theorem. The proof is relegated to Appendix B.9.

B Proofs

B.1 Proof of Proposition 2.1

Proof. For both hidden and transparent transaction costs, all market metrics are fully specified by gross demand and supply, as well as the trading volume $Q = D(P_B) = S(P_S)$. That is:

$$W = \int_{D^{-1}(Q)}^{2} D(x)dx + \int_{1}^{S^{-1}(Q)} S(x)dx \qquad (\textit{Total Welfare})$$
$$R = Q \cdot (D^{-1}(Q) - S^{-1}(Q)) \qquad (\textit{Revenue})$$
$$G^{real} = W + R \qquad (\textit{Realized Gains of Trade})$$
$$G^{gross} = \int_{P_{eq}}^{2} D(x)dx + \int_{1}^{P_{eq}} S(x)dx \qquad (\textit{Total Welfare})$$
$$L = \int_{P_{eq}}^{D^{-1}(Q)} (D(x) - Q)\, dx + \int_{S^{-1}(Q)}^{P_{eq}} (S(x) - Q)\, dx \qquad (\textit{Loss}).$$

For transparent transaction costs, the result is a direct consequence of Jantschgi et al. (2023, Theorem 10). Since the proof can be seamlessly adapted for hidden transaction costs, I will forgo further elaboration. For hidden transactions costs, the trading volume stands in a one to one bijection with the spread. Thus, any trading volume, and hence the associated market performance, can be implemented. For transparent transaction costs, the trading volume can be adjusted by a linear scaling. The fact that any trading volume — and hence any market performance associated with – is achievable, again follows from Jantschgi et al. (2023, Theorem 10). Thus, the same set of market performances are achievable with transparent and hidden transaction costs. □

B.2 Proof of Theorem 3.1

Theorem 3.1 is proven in several steps. In a first step, I show that due to the homogeneous arrival and memoryless departure, the market-clearing price and the pairs of buy and sell prices that balance trade are constant over time.

Lemma B.1. *For all $t \geq 1$, it holds that $P^{t,eq} = P^{0,eq}$. Moreover, for any buy price $P_b \in [P^{0,eq}, 2]$, consider the unique sell price $P_s \in [1, P^{0,eq}]$, such that $D^0(P_b) = S^0(P_s)$. Then, it holds for all trading rounds $t \geq 1$ that $D^t(P_b) = S^t(P_s)$.*

Proof of Lemma B.1. First, I prove by induction that the market clearing price is constant over time. For $t = 0$, the statement holds by definition. Next, assume that the statement holds for t, that is, $P^{t,eq} = P^{0,eq}$. $P^{t,eq} = P^{0,eq}$ is equivalent to $D^t(P^{0,eq}) = S^t(P^{0,eq})$. Note, that after clearing the market at prices P_b^t and P_s^t, it still holds that $Z_D^t(P^{0,eq}) = Z_S^t(P^{0,eq})$, see Appendix A.1. If $P \leq P_b^t$, it holds that $D^{t+1}(P) = \epsilon \cdot Z_D^t(P) + D^0(P)$. If $P \geq P_s^t$, it holds that $S^{t+1}(P) = \epsilon \cdot Z_S^t(P) + S^0(P)$. Note that $P_s^t \leq P^{0,eq} \leq P_b^t$. It holds that

$$D^{t+1}(P^{0,eq}) - S^{t+1}(P^{0,eq}) = \epsilon \cdot \left(Z_D^t(P^{0,eq}) - Z_S^t(P^{0,eq})\right) + \left(D^0(P^{0,eq}) - S^0(P^{0,eq})\right) \quad (6)$$

It holds by the induction statement that $Z_D^t(P^{0,eq}) - Z_S^t(P^{0,eq}) = 0$ and by definition of $P^{0,eq}$ that $D^0(P^{0,eq}) - S^0(P^{0,eq}) = 0$. Therefore, $D^{t+1}(P^{0,eq}) - S^{t+1}(P^{0,eq}) = 0$, which finally implies that $P^{t+1,eq} = P^{0,eq}$.

Second, I prove again by induction that the sequence of buy and sell prices that balance trade is constant over time. That is, for any buy price $P_b \in [P^{eq}, 1]$, consider the unique sell price P_s, such that $D^0(P_b) = S^0(P_s)$. Then, it holds for all trading rounds $t \geq 1$ that $D^t(P_b) = S^t(P_s)$. For $t = 0$, the statement holds by definition. Next, assume that the statement holds for t, that is, $D^t(P_b) = S^t(P_s)$. We show that this implies that also $D^{t+1}(P_b) = S^{t+1}(P_s)$. We need to consider two cases separately: (i) $P_b \in [P_b^t, 1]$ and (ii) $P_b \in [P^{eq}, P_b^t]$. For (i), it holds that $D^{t+1}(P_b) = D^0(P_b)$. Note that this implies that $P_s \in [0, P_s^t]$, and hence $S^{t+1}(P_s) = S^0(P_s)$. This implies that $D^{t+1}(P_b) = D^0(P_b) = S^0(P_s) = S^{t+1}(P_s)$. For (ii), it holds that $D^{t+1}(P_b) = \epsilon \cdot Z_D^t(P_b) + D^0(P_b)$ and $S^{t+1}(P_s) = \epsilon \cdot Z_S^t(P_s) + D^0(P_s)$. It thus holds that

$$D^{t+1}(P_b) - S^{t+1}(P_s) = \epsilon \cdot \left(Z_D^t(P_b) - Z_S^t(P_s)\right) + \left(D^0(P_b) - S^0(P_s)\right). \tag{7}$$

It follows from the assumption $D^t(P_b) = S^t(P_s)$ that $Z_D^t(P_b) = Z_S^t(P_s)$ holds as well, which implies that the first term in the equation above is equal to zero. Moreover, it follows from the definition of P_b and P_s, that the second term is also equal to zero. This finally implies that $D^{t+1}(P_b) - S^{t+1}(P_s) = 0$, which finishes the proof. □

Next, I show that at each point in time, the spread either decreases, or jumps back to the baseline spread σ^0. Moreover, a decreasing spread implies decreasing buy and increasing sell prices.

Lemma B.2 (Monotonicity)**.** *For all $t > 1$, it holds that either (i) $\sigma^t = \sigma^0$ or (ii) $\sigma^t < \sigma^{t-1}$, and thus $\sigma^t \leq \sigma^0$. If (ii) holds, then it must also hold that $P_b^t < P_b^{t-1}$, $P_s^t > P_s^{t-1}$ and $Q^t > Q^{t-1}$.*

Proof of Lemma B.2. I prove this Lemma via induction on the trading round t. For $t = 0$, the baseline spread σ^0 is optimal by definition and hence the statement holds for the base case. Next, consider that the hypothesis is true for the spread σ^{t-1} in round $t-1$. Let P_b^{t-1} and P_s^{t-1} be the buy and sell prices. For prices P above P_b^{t-1} it holds that $D^t(P) = D^0(P)$ and for prices P below P_s^{t-1} it holds that $S^t(P) = S^0(P)$. That is, because all buyers with order above P_b^{t-1} and all sellers with order below P_s^{t-1} that were in the market at time $t-1$ have been cleared. Hence, at these prices, there is no pent-up demand or supply in round t. It follows from Lemma B.1 and the strict monotonicity of demand and supply that a spread $\sigma^t \geq \sigma^{t-1}$ that clears the market corresponds to a buy price $P_b^t \geq P_b^{t-1}$ and a sell price $P_s^{t-1} \leq P_s^{t-1}$ in round t. Therefore, the revenue of such a spread in round t would yield the revenue $(P_b^t - P_s^t) \cdot D^t(P_b^t) = (P_b^t - P_s^t) \cdot D^0(P_b^t)$. However, for all such spreads, the baseline spread σ^0 maximizes the revenue at $(P_b^0 - P_s^0) \cdot D^0(P_b^0)$. Thus, it must hold that $\sigma^t = \sigma^0$ or $\sigma^t < \sigma^{t-1}$, which proves the induction step. Next, consider a spread $\sigma^t < \sigma^{t-1}$. It follows again from Lemma B.1 and the strict monotonicity of demand and supply that a spread $\sigma^t < \sigma^{t-1}$ that clears the market corresponds to a buy price $P_b^t < P_b^{t-1}$ and a sell price $P_s^{t-1} > P_s^{t-1}$ in round t. Finally, it follows from the evolution of demand and supply, that $P_b^t < P_b^{t-1}$ implies $Q^t = D^t(Pt) > D^{t-1}(P-1) = Q^{t+1}$. □

Lemma B.2 implies that the baseline spread σ^0 is an upper bound for the evolution of bid-ask spreads. No matter the history, the market platform at time t will never choose a bigger spread.

Corollary B.3 (Upper bound on the bid-ask spread)**.** *For all $t \geq 1$, it holds that $\sigma^t \leq \sigma^0$.*

Next, I show that there also exists a lower bound on the bid-ask spread, that is, the sequence of realized spreads cannot reach values, that are arbitrarily close to zero.

Lemma B.4 (Lower bound on the bid-ask spread)**.** *There exists a spread $\sigma^* > 0$, such that for all $t > 1$, it holds that $\sigma^t \geq \sigma^*$.*

Proof of Lemma B.4. Due to a non-zero departure rate, built-up demand and supply does not grow to infinity. It follows from Lemma A.1, that in any round t it holds that $D^t(P) \leq \frac{1-\epsilon^{t+1}}{1-\epsilon} D^0(P) < \frac{1}{1-\epsilon}$ and $S^t(P) \leq \frac{1-\epsilon^{t+1}}{1-\epsilon} S^0(P) < \frac{1}{1-\epsilon}$. Consider time t with corresponding demand $D^t(\cdot)$ and supply $S^t(\cdot)$. For any feasible spread σ, let P_b and P_s be the unique buy and sell prices that correspond to the spread σ; recall that by Lemma B.1 this does not depend on time t. It follows from Lemma B.4 that we can restrict our attention to spread $\sigma \leq \sigma^0$, such that $P_b \leq P_b^0$ and $P_s \geq P_s^0$. For such a spread, the revenue for the market platform is thus equal to $\sigma \cdot D^t(P_b)$. It follows from the evolution of demand and supply (see Lemma A.1), that

$$D^t(P_b) \leq D^0(P_b) + \left(D^0(P_b) - D^0(P_b^0)\right) \frac{1-\epsilon^t}{1-\epsilon}.$$

This implies the following crude upper bound on the trading volume at buy price P_b: $D^t(P_b) < \frac{1}{1-\epsilon} D^0(P_b) < \frac{1}{1-\epsilon}$. The first inequality corresponds to a market that has never been cleared for an infinite period of time, that is, all demand up to price P_b has built up. The second inequality is due to the fact that a unit mass of traders arrives in each round. Now, consider the baseline spread σ^0, which leads to a revenue equal to $R^0 = \sigma^0 \cdot Q^0 = \sigma^0 \cdot D^0(P_b^0)$. Consider σ^* such that $\sigma^* \cdot \frac{1}{1-\epsilon} = R^0$. At no time t, the market platform would choose a spread below σ^*, as they could strictly improve their revenue by charging the spread σ^0. □

The next Lemma shows that the baseline spread is repeated infinitely often, but that also for infinitely many times, a strictly smaller spread is realized, proving the existence of cycles:

Lemma B.5 (Cycles)**.** *There exist infinitely many t, such that $\sigma^t = \sigma^0$. If the departure rate $1-\epsilon$ is sufficiently low, there exist infinitely many t', such that $\sigma^t \neq \sigma^0$.*

Proof of Lemma B.5. It follows from Lemma B.2 that for all $t > 1$, it holds that either (i) $\sigma^t = \sigma^0$ or (ii) $\sigma^t < \sigma^{t-1}$. First, assume that for all $t \geq 1$, it holds that $\sigma^t = \sigma^0$. The revenue in each round is thus equal to $\sigma^0 \cdot Q^0$, where $Q^0 = D^0(P_b^0)$ denotes the baseline trading volume. Consider a small $\delta > 0$. We will show that for sufficiently small δ, sufficiently large t and sufficiently large ϵ, the spread $\sigma - \delta$ will strictly increase the revenue. This then implies that it is not optimal to have $\sigma^t = \sigma^0$ for all t. Consider now that up until round t, the market was always cleared at the baseline spread σ. Clearing the market now at the spread $\sigma - \delta$, denote by P_b and P_s the corresponding buy and sell prices in round t. We need to estimate the corresponding trading volume $D^t(P_b)$. As we clear at a smaller spread, we definitely clear all traders, that we would have cleared at the baseline spread with a mass equal to Q^0. Moreover, decreasing the spread by δ means that it must either hold that $P_b^0 - P_b \geq \frac{\delta}{2}$ or $P_s - P_s \geq \frac{\delta}{2}$. That is, either buy or sell price must have changed by at least $\frac{\delta}{2}$. By assumption, the derivative of demand and supply are strictly bounded away from zero, that is, there exists $\gamma > 0$, such that $D'(P) \leq -\gamma$ and $S'(P) \geq \gamma$ for all $P \in [0,1]$. Thus, it holds that only considering the baseline demand and supply, the trading volume must have increased by at least $\frac{\delta}{2}\gamma$, as the buy or sell price changed by at least $\frac{\delta}{2}$. However, as excess demand and supply has built up for t rounds, as the market was always cleared at the baseline spread, the following inequality holds for the trading volume: $Q^t \geq Q^0 + \frac{\delta}{2}\gamma\frac{1-\epsilon^t}{1-\epsilon}$. Thus, in order for the baseline spread to still be optimal in round t, it must hold that

$$\sigma^0 Q^0 > (\sigma^0 - \delta) \cdot \left(Q^0 + \delta\gamma\frac{1-\epsilon^t}{1-\epsilon}\right). \tag{8}$$

Reordering the terms yields the inequality $Q^0 > (\sigma - \delta)\frac{\gamma}{2}\frac{1-\epsilon^t}{1-\epsilon}$. Note that $Q^0 < 1$ and $\sigma^0 > 0$ are constants given by the model. For sufficiently small $\delta > 0$, for sufficiently large t and for ϵ sufficiently close to 1, this yields a contradiction. Thus, it cannot be that for all $t \geq 1$ it holds that $\sigma^t = \sigma^0$.

Next, assume by contradiction that σ^0 is not realized infinitely often. Then, Lemma B.2 implies that there exists a $T > 0$, such that for all $t \geq T$, σ^t is a strictly decreasing sequence. Without loss of generality, assume that $T = 1$. It follows from Lemma B.4, that there exists a lower bound σ^*. The sequence $(\sigma^t)_{t\geq 1}$ must therefore converge to a limit $\sigma^\infty > \sigma^* > 0$. Note that by Lemma B.2 the sequence of buy prices $(P_b^t)_{t\geq 1}$ is strictly decreasing with limit P_∞^b and the sequence of sell prices $(P_s^t)_{t\geq 1}$ is strictly increasing with limit P_∞^s. Thus, for all $\delta > 0$, there exists a time t', such that for all $t \geq t'$, it holds that $P_b^t - P_b^{t+1} \leq \delta$ and $P_s^{t+1} - P_s^t \leq \delta$. Clearing a spread of σ^{t+1} has two effects: It clears the newly arriving buyers and sellers, and it clears some additional pent-up demand and supply, that was not cleared in the previous round. Note that for a strictly decreasing sequence of spreads, in round $t+1$, you only clear excess demand and supply from the price intervals $[P_b^{t+1}, P_b^t]$ and $[P_s^t, P_s^{t+1}]$. The slope of baseline demand is lower bounded and the slope of baseline supply is upper bounded, and the difference between prices in rounds t and $t+1$ is less than δ. Moreover, the difference in slope between baseline demand and demand in round t is upper bounded by the constant $\frac{1}{1-\epsilon}$. Thus, the amount of excess demand and supply that is cleared from round t to round $t+1$ is upper bounded by $\delta\gamma\frac{1}{1-\epsilon}$. Hence, the revenue in round $t+1$ is strictly upper bounded by $\sigma^{t+1} \cdot D^0(P_b^{t+1}) + \sigma^{t+1} \cdot \delta\gamma\frac{1}{1-\epsilon}$. Letting t tend to infinity, the revenue thus approaches $\sigma^\infty \cdot D^0(P_\infty^b)$. However, the spread σ^∞ is not optimal for baseline demand and supply, contradicting the optimality of a strictly decreasing sequence of spreads $(\sigma^t)_{t\geq 1}$. Thus, the spread must jump back to the baseline spread infinitely often, proving the claim. □

B.3 Proof of Theorem 3.2

Proof. I first show that for all $t \geq t_{n-1}$, if $\sigma^t \geq \sigma^{t_{n-1}}$, then $\sigma^t = \sigma^{t-t_{n-1}}$. Suppose that the statement does not hold for some t. Then, there exists a first time, where C_n deviates from $\bigcup_{i=1}^{n-1} C_i$ and this deviation is not a new global minimum. Suppose that this deviation occurs at time t', and suppose that $t' = t_{n-1} + k$ for some $k > 0$. Let $\sigma^{t'}$ be the associated spread. By assumption $\sigma^{t'} \neq \sigma^{t'-t_{n-1}} = \sigma^k$, and $\sigma^{t'} \geq \sigma^{t_{n-1}}$. Note that at time $t_{n-1} + 1$, the cycle C_n begins for the first time. At time t_{n-1}, the spread $\sigma^{t_{n-1}}$ was a new global minimum. Let $P_b^{t_{n-1}}$ be the associated buy price and $P_s^{t_{n-1}}$ the associated sell price. At time $t_{n-1} + 1$, there is no excess demand for prices above $P_b^{t_{n-1}}$ and no excess supply for prices below $P_s^{t_{n-1}}$. By assumption, t' was the first deviation. That is, in the the two time intervals $[1, k]$, and $[t_{n-1} + 1, t']$, the market is cleared at exactly the same spreads, and all of these spreads are greater or equal than $\sigma^{t_{n-1}}$. Thus, the only difference of built-up excess demand and supply is for buy prices below $P_b^{t_{n-1}}$ and sell prices above $P_s^{t_{n-1}}$. However, until time t', such prices are never realized. This implies that for prices $P \geq P_b^{t_{n-1}}$, it holds that $D^{t'}(P) = D^k(P)$. Similarly, for prices $P \leq P_s^{t_{n-1}}$, it holds that $S^{t'}(P) = S^k(P)$. Note that both spread $\sigma^{t'}$ and σ^k are greater or equal than $\sigma^{t_{n-1}}$. Thus, the buy prices associated both with spread $\sigma^{t'}$ and σ^k are greater or equal to $P_b^{t_{n-1}}$. Similarly, the sell prices associated both with spread $\sigma^{t'}$ and σ^k are less or equal to $P_s^{t_{n-1}}$. Hence, $R^{t'}(\sigma^{t'}) = R^{t'}(\sigma^k)$ and $R^k(\sigma^{t'}) = R^k(\sigma^k)$. If $\sigma^{t'}$ were optimal at time t', it woulds also be optimal at time k. However, at this time, the spread σ^k was optimal. This now yields a contradiction. Thus, if at any point in the cycle C_n the spread is different from $\bigcup_{i=1}^{n-1} C_i$, it must be a new global minimum.

Moreover, this implies that each cycle C_n is finite. Consider the spread $\sigma^{2t_{n-1}} \in C_n$. If there was no strict global minimum up to this time, then the statement above implies that $\sigma^{2t_{n-1}} = \sigma^{t_{n-1}}$. Thus, by definition of t_n, the cycle ends, once the baseline spread σ^0 is realized again. It follows from Theorem 3.1, that for all $t \geq 2t_{n-1}$ the spread either strictly decreases or jumps back up to σ^0. However, as σ^0 is realized infinitely often, it cannot be that the sequence is strictly decreasing forever. Thus, t_n is finite, which shows that the cycle C_n has finite length. □

B.4 Proof of Theorem 4.1

Proof. In order to shows that sequence of buy and sell prices is constant over time, recall that $P_b^t = P^{t,eq} + \Phi_B(P^{t,eq})$ and $P_s^t = P^{t,eq} - \Phi_S(P^{t,eq})$. Thus, buy and sell prices are completely determined by the market-clearing price $P^{t,eq}$ with respect to revealed demand and supply. It is thus sufficient to show the following:

Lemma B.6. *For all trading rounds $t \geq 1$, it holds that $P^{t,eq} = P^{0,eq}$.*

Note that Lemma B.6 seems identical to Lemma B.1. However, the difference is that for transparent transaction costs, the market-clearing price corresponds to revealed net demand and supply instead of the true demand and supply. In the case of net demand and supply, the proof is even simpler. In the first round, there is no built-up excess demand and supply. Thus, $P^{0,eq}$ is the market-clearing price with respect to $D^{0,net}(\cdot)$ and $S^{0,net}(\cdot)$, implying that

$P^{1,eq} = P^{0,eq}$. Recall that for transparent transaction costs, the allocation is determined by the market-clearing price, the transaction costs are added ex post. That is, all buy orders with revealed value above $P^{0,eq}$ and all sell orders with revealed value above $P^{0,eq}$ are cleared. Thus, excess demand only builds up below $P^{0,eq}$ and excess supply only builds up above $P^{0,eq}$, see Appendix A for a more formal discussion. For all $P \geq P^{0,eq}$, it holds that $D^2(P) = D^0(P)$ and for all $P \leq P^{0,eq}$, it holds that $S^2(P) = S^0(P)$. Thus, in the second clearing round, it holds that $D^2(P^{0,eq}) = S^2(P^{0,eq})$. This proves that in the second clearing round, $P^{0,eq}$ is still the market clearing price, which implies $P^{2,eq} = P^{0,eq}$. Lemma B.6 now follows from a straightforward inductive argument. □

B.5 Proof of Theorem 5.1

Proof. First, I show that for every $\alpha \in \mathbb{R}$, it holds that $\mathbb{E}[u_\alpha(\sigma^H)] > u_\alpha(\sigma^0)$, and $\mathbb{E}[u_\alpha(\sigma^H)] < u_\alpha(0)$ It follows from the law of total expectation that $\mathbb{E}_i[u_\alpha(\sigma^H)] = \sum_{j\geq 0} u_\alpha(\sigma^j)\mathbb{P}_i[\sigma^H = \sigma^j]$. It follows from Theorem 3.1 that for all $j \geq 0$, it holds that $\sigma^j \leq \sigma^0$. Because the utility is strictly decreasing in the spread, it holds that $u_\alpha(\sigma^0) < u_\alpha(\sigma^j)$ for all $j \geq 0$. Note that $\sum_{j\geq 0} \mathbb{P}_i[\sigma^H = \sigma^j] = 1$. Hence, $\mathbb{E}[u_\alpha(\sigma^H)] \geq \sum_{j\geq 0} u_\alpha(\sigma^0)\mathbb{P}_i[\sigma^H = \sigma^j] = u_\alpha(\sigma^0)$. Thus, a trader weakly prefers the platform with hidden transaction costs. However, by assumption, there exists at least one j, such that $\sigma^j \neq \sigma^0$ and $\mathbb{P}_i[\sigma^H = \sigma^j] > 0$. Because the utility is strictly decreasing, this implies that the inequality is actually strict, that is $\mathbb{E}[u_\alpha(\sigma^H)] > u_i(\sigma^0)$. Thus, a trader strictly prefers a market platform with hidden transaction costs and associated spread σ^H over a market with transparent transaction costs and spread σ^T, if $\sigma^T = \sigma^0$. The argument for $\mathbb{E}[u_\alpha(\sigma^H)] < u_\alpha(0)$ is analogous.

Existence: For a fixed level of risk-aversion α, consider the map $\sigma^T \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$ for $\sigma^T \in [0, \sigma^0]$. This map is continuous and strictly decreasing, because $u_\alpha(\cdot)$ is continuous and strictly decreasing. Moreover, it follows from the previous observation above, that evaluated at the point σ^0, the map is strictly negative, while evaluated at the point 0, the map is strictly positive. Thus, the Intermediate Value Theorem implies the existence of a unique zero point, which corresponds to the equilibrium.

Sub-monopoly pricing: It follows directly from the observation $\mathbb{E}[u_\alpha(\sigma^H)] > u_\alpha(\sigma^0)$ and the strict monotonicity of $u_\alpha(\cdot)$, that in equilibrium, it must hold that $\sigma^T < \sigma^0$.

Comparative statics: For a fixed transparent spread σ^T, consider the map $\alpha \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$. This map is continuous and strictly increasing. To see that the map is continuous, note that $u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)] = u_\alpha(\sigma^T) - \sum_{j\geq 0} u_\alpha(\sigma^j)\mathbb{P}[\sigma^H = \sigma^j]$ holds and the map $\alpha \mapsto u_\alpha(\cdot)$ is continuous. The strict monotonicity follows from the fact that $u_\alpha(\cdot)$ has constant relative risk aversion (CRRA). This implies that as α increases, the function $u_\alpha(\cdot)$ becomes more and more concave. Thus, a trader will prefer the certainty of outcome σ^T over the random nature of the outcome σ^H, proving that the map $\alpha \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$ is indeed strictly increasing. As the equilibrium corresponds to the unique zero of this function, note that for a fixed transparent spread σ^T, an increase in α leads to a positive value of the function. As, for fixed α, the function is continuous and strictly decreasing in α, in order to have a zero point at a higher α, the transparent spread must also increase. Hence, in equilibrium, the transparent spread is increasing as a function of α.

Efficiency Crossover. First, I will argue that if I let α go to ∞, in equilibrium, the transparent spread converges to σ^0. Moreover, if I let α go to $-\infty$, in equilibrium, the transparent spread converges to σ^0. To see this, consider the equilibrium equality $\mathbb{E}[u_\alpha(\sigma^H)] = u_\alpha(\sigma^T)$. Recall that $u_\alpha(\sigma) = \frac{1-exp(-\alpha\cdot(1-\sigma))}{\alpha}$ for $\alpha \neq 1$. Recall that traders have beliefs about the realization of σ^H, more specifically, full support beliefs over a set $\sigma^{H1} = \sigma^0 > \sigma_1^H ... > \sigma_n^H > 0$ with corresponding probabilities $p_1, ..., p_n > 0$. Hence, we have that

$$\mathbb{E}[u_\alpha(\sigma^H)] = \sum_{i=1}^{n} p_i u_\alpha \sigma_i^H = \sum_{i=1}^{n} p_i \cdot \frac{1 - exp(-\alpha \cdot (1 - \sigma_i^H))}{\alpha}. \tag{9}$$

Using linearity of expectation, we get that

$$\mathbb{E}[u_\alpha(\sigma^H)] = \frac{1}{\alpha}\left(1 - \sum_{i=1}^{n} p_i \cdot exp(-\alpha \cdot (1 - \sigma_i^H))\right). \tag{10}$$

Hence, the equilibrium equality implies that

$$\frac{1 - exp(-\alpha \cdot (1 - \sigma^T))}{\alpha} = \frac{1}{\alpha}\left(1 - \sum_{i=1}^{n} p_i \cdot exp(-\alpha \cdot (1 - \sigma_i^H))\right). \tag{11}$$

Using basic algebra, this yields the following analytical expression for the transparent spread σ^T:

$$\sigma^T = 2 + \frac{1}{\alpha} \ln\left(\sum_{i=1}^{n} p_i \cdot exp(-\alpha \cdot (1 - \sigma_i^H))\right). \tag{12}$$

Now, let's analyze the behavior of σ^T as $\alpha \to \infty$ and $\alpha \to -\infty$.

First, I analyze $\alpha \to \infty$. As $\alpha \to \infty$, the terms $\exp(-\alpha \cdot (1 - \sigma_i^H))$ will be dominated by the largest σ_i^H, i.e., the smallest $1 - \sigma_i^H$. Given the ordering $\sigma_1^H = \sigma^0 > \sigma_2^H > \cdots > \sigma_n^H$, the dominant term will be the one corresponding to the baseline spread $\exp(-\alpha \cdot (1 - \sigma^0))$. Therefore, for large α,

$$\sum_{i=1}^{n} p_i \exp(-\alpha \cdot (1 - \sigma_i^H)) \approx p_1 \exp(-\alpha \cdot (1 - \sigma^0)). \tag{13}$$

Thus, for large α, it holds that

$$\sigma^T \approx 1 + \frac{1}{\alpha} \ln\left(p_1 \exp(-\alpha \cdot (1 - \sigma^0))\right). \tag{14}$$

Simplifying the logarithm yields

$$\ln\left(p_1 \exp(-\alpha \cdot (1 - \sigma^0))\right) = \ln(p_1) - \alpha \cdot (1 - \sigma^0). \tag{15}$$

Hence, it holds that

$$\sigma^T \approx 1 + \frac{1}{\alpha}(\ln(p_1) - \alpha \cdot (1 - \sigma^0)). \tag{16}$$

As $\alpha \to \infty$, the term $\frac{\ln(p_1)}{\alpha}$ converges to 0, leaving: $\sigma^T \approx 1 - (1 - \sigma^0) = \sigma^0$ Therefore, as $\alpha \to \infty$, $\sigma^T \to \sigma^0$. For $\alpha \to -\infty$, the other extreme term σ_n^H dominates. Thus, by the same proof, we can conclude that σ^T converges to σ_n^H. In summary: As $\alpha \to \infty$, the transparent spread σ^T, in equilibrium, converges to σ_1^H. As $\alpha \to -\infty$, the transparent spread σ^T, in equilibrium, converges to σ_n^H.

Next, I prove that the long-run average efficiency of the market with transparent transaction costs is continuous and strictly decreasing as a $\alpha \mapsto E^\infty(\sigma^T)$. First, as the transparent spread remains constant over time, note that the long-run average efficiency is equal to the per-round efficiency. Moreover, it was just shown that a strict increase in the risk aversion, in equilibrium, leads to a strict increase in the transparent spread. However, a strictly bigger spread leads to strictly smaller realized gains of trade — equivalently a strictly bigger loss — and hence strictly smaller efficiency. A formal argument can be found in Jantschgi et al. (2023). Intuitively, see the right-hand side of Figure 1: For transparent transaction costs, the loss is equal to the triangle between true demand and supply and the revenue rectangle. A bigger spread scales down the rectangle, leading to a strictly larger loss-triangle. The continuity follows from the fact that a slight increase in the spread continuously increases the loss-triangle. Thus, the following observation holds:

Observation B.7. For sufficiently large α, it thus holds that $E^\infty(\sigma^T)$ is close to $E^\infty(\sigma^0)$, and for sufficiently small α, $E^\infty(\sigma^T)$ is close to $E^\infty(\sigma^\infty)$.

Next, I analyze the long-run average efficiency of the market with hidden transaction costs, leading to the cyclical sequence of spreads. $\sigma^H = (\sigma^{Ht})_{t\geq 1}$. Let σ_∞^H be the spread that converges to the global minimum — this is well defined by Theorems 3.1 and 3.2. Hence, we study $E^\infty(\sigma^H) = \lim_{T\to\infty} \frac{\sum_{t=1}^T G^{real}(\sigma_t^H)}{T \cdot G^{gross}}$.

Observation B.8. It holds that $E^\infty(\sigma^0) < E^\infty(\sigma^H) < E^\infty(\sigma_\infty^H)$.

I start by showing that $E^\infty(\sigma^0) < E^\infty(\sigma^H)$. I define the following sequence of stopping times: Let T_1 be the first $t \geq 0$, such that $\sigma^{H_t} \neq \sigma^0$, that is, the first time the spread is an improvement to the baseline spread σ^0. Iteratively, let T_n be the first time after T_{n-1}, when the spread is again an improvement to the baseline spread. It follows from Theorem 3.1 that for any n, T_n is finite, as the realized spread is infinitely often not equal to the baseline spread. However, the predictable price cycles from Theorem 3.2, allow to tighten that statement. Recall that at the end of every price cycle C_n, a global minimum occurs. However, Theorem 3.2 implies that n'th price cycle C_n repeats $\bigcup_{i=1}^{n-1} C_i$, until it is broken by a new global minimum. That is, the cycles repeat, until a new global minimum spread is realized. This however implies that the time between $T_n - T_{n-1}$ is bounded by some constant that does not depend on n. Hence, the time between spreads that are not equal to the baseline spread does not diverge, but is bounded by some finite number k. We rewrite $E^{\infty}(\sigma^H) = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} G^{real}(\sigma_t^H)}{T \cdot G^{gross}}$ as follows:

$$\lim_{T\to\infty} \frac{\sum_{t=1}^{T} G^{real}(\sigma_t^H)}{T \cdot G^{gross}} = \lim_{n\to\infty} \frac{\sum_{i=1}^{n} \sum_{t=T_i}^{T_{i+1}} G^{real}(\sigma_t^H)}{\sum_{i=1}^{n} (T_{i+1} - T_i) G^{gross}}. \tag{17}$$

Note that in every interval $[T_i, T_{i+1}]$, there is one spread that is strictly less than σ^0. As every new realized spread is a global minimum (Theorem 3.2), it holds that this spread is also bounded from above by some spread $\sigma^* < \sigma^0$. Note that no matter the build-up of excess demand and supply, clearing at a spread σ^* instead of σ^0 improves the realized gains of trade by at least some fixed $\delta > 0$ (just accounting for the additional executed trades from newly incoming demand and supply). Thus, for every i it holds that

$$\sum_{t=T_i}^{T_{i+1}} G^{real}(\sigma_t^H) \geq \sum_{t=T_i}^{T_{i+1}} G^{real}(\sigma^0) + \delta. \tag{18}$$

Combining the two equations from above yields that $E^{\infty}(\sigma^H) > E^{\infty}(\sigma^0)$. To show that $E^{\infty}(\sigma^H) < E^{\infty}(\sigma^{\infty})$, I use a similar argument. Note that every time, when the market clears at σ^{∞} instead of σ^0, there is an improvement in the realized gains of trade by at least $\delta > 0$. Next, define a similar sequence of stopping times for the events, when σ^H is equal to the baseline spread σ^0. It again follows from *Theorems* 3.1*and* 3.2 that these events not only happen infinitely often, but that the time between these events is also upper bounded by some uniform constant. The analogous argument to above — rewriting the sum over the intervals of these stopping times and at least one δ improvement per interval – yields that $E^{\infty}(\sigma^H) < E^{\infty}(\sigma^{\infty})$.

Combining Observation B.7 and Observation B.8 finishes the proof.

□

B.6 Proof of Theorem 6.2

Proof. Let σ^* be the smallest spread, such that clearing the baseline demand and supply yields a revenue of at least the revenue threshold R. If any smaller spread is realized at some point, market failure occurs. Thus, it suffices to prove that for every $k \geq 1$, it holds that $\sigma^{inf,k} < \sigma^*$ for a sufficiently low departure rate. Suppose that for level k best responses, the new baseline spread for the price cycles is $\sigma^{sup,k}$ with revenue $R^{sup,k}$ and let $\sigma^{inf,k}$ be the new infimum. I show that for a sufficiently low departure rate, it must hold that $\sigma^{inf,k} < \sigma^*$. Suppose not. Let $\delta > 0$ denote the mass of traders in the level k-baseline market, whose order is executed at spread σ^*, but not at spread $\sigma^{inf,k}$. The orders of these traders are thus never executed. At time t, mass of traders with such orders is $\frac{1-\epsilon^t}{1-\epsilon} \cdot \delta$. If the market were to clear at spread $\sigma^{inf,k}$, the revenue from these traders alone would be $\frac{1-\epsilon^t}{1-\epsilon} \cdot \delta \cdot \sigma^{inf,k}$. For sufficiently large t and sufficiently small $\epsilon > 0$, this revenue is greater than the fixed level

k baseline revenue at $\sigma^{sup,k}$. Because this spread is realized infinitely often, there exists a large t, where this revenue is realized. However, clearing at $\sigma^{inf,k}$ would improve revenue. This is a contradiction. Hence, for any k, for a sufficiently low departure rate, it must hold $\sigma^{inf,k} < \sigma^*$, which means that there exist times, where a spread strictly smaller than σ^* is realized. However, for level $k+1$ best responses, the new baseline is smaller or equal to $\sigma^{inf,k}$ and thus strictly smaller than σ^*. Because σ^* was the smallest spread, for which the market without excess was profitable, this means that in round $k+1$, market failure occurs. □

B.7 Proof of Lemma A.1

Proof. I will prove the statement for the demand functions, the proofs for the supply functions are analogous. For $P_b' > P_b$, we are interested in the evolution of $D^t(P_b) - D^t(P_b')$. The pent-up demand in the range $[P_b, P_b']$ evolves as follows: First, every time step a new amount of $D^0(P_b) - D^0(P_b')$ is added to the pent-up demand. Whenever some P_b^t is larger than P_b, then all of this pent-up demand is supplied and the pent-up demand in that interval for time $t+1$ is zero. When P_b^t is larger than P_b', all of the pent-up demand remains and carries over by the factor ϵ into the next. Whenever $P_b < P_b^t < P_b'$, then some of this demand is supplied and some remains and carries over by the factor ϵ into the next round. More formally:

1. When $P_b^{t-1} < P_b < P_b'$, there is no pent-up demand $Z_D^{t-1}(P_b) = Z_D^{t-1}(P_b') = 0$ and thus
$$D^t(P_b) - D^t(P_b') = D^0(P_b) - D^0(P_b'). \tag{19}$$
2. When $P_b^{t-1} > P_b' > P_b$, we have that $Z_D^{t-1}(P_b) = D^{t-1}(P_b) - Q^{t-1}$ and $Z_D^{t-1}(P_b') = D^{t-1}(P_b') - Q^{t-1}$. This implies $Z_D^{t-1}(P_b) - Z_D^{t-1}(P_b') = D^{t-1}(P_b) - D^{t-1}(P_b')$ and thus
$$D^t(P_b) - D^t(P_b') = \epsilon \cdot \left(D^{t-1}(P_b) - D^{t-1}(P_b')\right) + \left(D^0(P_b) - D^0(P_b')\right). \tag{20}$$
3. When $P_b \le P_b^{t-1} \le P_b'$, it holds that $Z_D^{t-1}(P_b') = 0$, while $Z_D^{t-1}(P_b) = D^{t-1}(P^b_) - Q^{t-1}$. Because $D^{t-1}(P_b') \le Q^{t-1} = D^{t-1}(P_b^{t-1})$, it follows that
$$D^{t-1}(P_b) - D^{t-1}(P_b') \ge Z_D^{t-1}(P_b) - Z_D^{t-1}(P_b). \tag{21}$$

   Note that the last equation holds in all three cases.

To prove (1), we observe that

$$\begin{aligned} D^t(P_b) - D^t(P_b') &= \epsilon \cdot Z_D^{t-1}(P_b) + D^0(P_b) - \left(\epsilon \cdot Z_D^{t-1}(P_b') + D^0(P_b')\right) \\ &= \epsilon \cdot \left(Z_D^{t-1}(P_b) - Z_D^{t-1}(P_b')\right) + \left(D^0(P_b) - D^0(P_b')\right) \\ &\le \epsilon \cdot \left(D^{t-1}(P_b) - D^{t-1}(P_b')\right) + \left(D^0(P_b) - D^0(P_b')\right) \end{aligned} \tag{22}$$

Using the same reasoning, we get that

$$D^{t-1}(P_b) - D^{t-1}(P_b') \le \epsilon \cdot \left(D^{t-2}(P_b) - D^{t-2}(P_b')\right) + \left(D^0(P_b) - D^0(P_b')\right). \tag{23}$$

Combining these two inequalities gives

$$D^t(P_b) - D^t(P_b^{'}) \leq \epsilon^2 \cdot \Big(D^{t-2}(P_b) - D^{t-2}(P_b^{'})\Big) + (1+\epsilon) \cdot \Big(D^0(P_b) - D^0(P_b^{'})\Big). \tag{24}$$

Applying this reasoning inductively, we arrive at the final inequality

$$\begin{aligned} D^t(P_b) - D^t(P_b^{'}) &\leq (1+\epsilon+\epsilon^2+\ldots+\epsilon^t) \cdot \Big(D^0(P_b) - D^0(P_b^{'})\Big) \\ &= \frac{1-\epsilon^{t+1}}{1-\epsilon}\Big(D^0(P_b) - D^0(P_b^{'})\Big) \leq \frac{1}{1-\epsilon}\Big(D^0(P_b) - D^0(P_b^{'})\Big). \end{aligned} \tag{25}$$

The proof for (2) is analogous, if the iterative reasoning from above is stopped after the first $t-T$ steps. For (3), we note that when $P_b^{t-1} \geq P_b^{'} \geq P_b$, it holds that $Z_D^{t-1}(P_b) = D^{t-1}(P_b) - Q^{t-1}$ and $Z_D^{t-1}(P_b^{'}) = D^{t-1}(P_b^{'}) - Q^{t-1}$. Hence, $D^t(P_b) - D^t(P_b) = \epsilon \cdot (Dt-1P_b - Dt-1P_b) - \big(D^0(P_b) - D^0(P_b^{'})\big)$, and we get equality throughout the induction that is used to prove (2). (4) and (5) follow as applications from (2) and (3) respectively, because it holds that $Z_D^t(P_b) = Z_D^t(P_b) = 0$, if $P_b^t \leq P_b \leq P_b^{'}$ (or $T = 0$). □

B.8 Proof of Theorem A.2

Proof. First, I prove the following auxiliary Lemma: The price cycles from myopic revenue maximization improve the long-run average revenue compared to clearing at a constant spread. Note that, while myopic revenue maximization offer a weak improvement in every round, this does not directly yield the result. It is neccessary to show that these revenue improvements happen sufficiently often to have an impact on the average revenue.

Lemma B.9. *Let $R^\infty(\sigma^c)$ be the long-run average revenue of clearing at constant spread σ^c, and let $R^\infty(\sigma^{myopic})$ be the long-run average revenue of the price cycles from myopic revenue maximization in Theorems 3.1 and 3.2. There exists $\delta > 0$, such that for all $\sigma^c \geq 0$ it holds that $R^\infty(\sigma^c) \leq R^\infty(\sigma^{myopic}) - \delta$.*

Proof. I will only prove this for the optimal baseline spread σ^0. Any other spread has weakly smaller revenue per round, and thus also weakly smaller long-run average revenue: $R^\infty(\sigma^c) \leq R^\infty(\sigma^0)$.

Let T_1 be the first $t \geq 0$, such that $\sigma_t^H \neq \sigma^0$, that is, the first time the spread is an improvement to the baseline spread σ^0. Iteratively, let T_n be the first time after T_{n-1}, when the spread is again an improvement to the baseline spread. It follows from Theorem 3.1 that for any n, T_n is finite, as the realized spread is infinitely often not equal to the baseline spread. However, the predictable price cycles from Theorem 3.2 allow to tighten that statement. Recall that at the end of every price cycle C_n, a global minimum occurs. However, Theorem 3.2 implies that the n'th price cycle C_n repeats $\bigcup_{i=1}^{n-1} C_i$, until it is broken by a new global minimum. That is, the cycles repeat until a new global minimum spread is realized. This implies that the time between $T_n - T_{n-1}$ is bounded by some constant that does not depend on n. Hence, the time between spreads that are not equal to the baseline spread does not diverge, but is bounded by some finite number k.

Rewrite the long-run average revenue by splitting it up over the intervals $[T_i, T_{i+1}]$:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} R^t(\sigma^c) = \lim_{n\to\infty} \frac{\sum_{i=1}^{n} \sum_{t=T_i}^{T_{i+1}} R^t(\sigma^c)}{\sum_{i=1}^{n} (T_{i+1} - T_i)}. \tag{26}$$

It follows from Theorems 3.1 and 3.2 that the intervals T_i are finite and that in each interval $[T_i, T_{i+1}]$, there is at least one time point where the revenue is strictly greater than the revenue from the baseline spread σ^0. Moreover, it holds that there exists a lower bound $\delta > 0$ on the revenue improvements that does not depend on i.

In every interval $[T_i, T_{i+1}]$, we have:

$$\sum_{t=T_i}^{T_{i+1}} R^t(\sigma^0) \le \sum_{t=T_i}^{T_{i+1}} R^t(\sigma^{myopic}) - \delta. \tag{27}$$

Since the length of the intervals $[T_i, T_{i+1}]$ are bounded by some constant k that does not depend on i, we can write:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} R^t(\sigma^0) \le \lim_{n\to\infty} \frac{1}{\sum_{i=1}^{n} (T_{i+1} - T_i)} \sum_{i=1}^{n} \left(\sum_{t=T_i}^{T_{i+1}} R^t(\sigma^{myopic}) - \delta \right). \tag{28}$$

This simplifies to:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} R^t(\sigma^0) \le \lim_{n\to\infty} \left(R^\infty(\sigma^{myopic}) - \frac{\delta}{k} \right). \tag{29}$$

Hence, for any constant spread $\sigma^c \ge 0$:

$$R^\infty(\sigma^c) \le R^\infty(\sigma^{myopic}) - \frac{\delta}{k}. \tag{30}$$

This completes the proof. □

To show that no eventually monotone sequence of spreads maximizes the long-run average revenue, I will construct an oscillatory sequence of spreads that strictly improves the long-run average revenue.

For notational simplicity, I will only consider monotone sequences of spreads, that is, eventually monotone sequences with $t = 1$. If the sequence or spreads starts with an oscillatory part, the constructed improving sequence will simply choose the same sequence, up until the round, when the sequence becomes weakly monotone. This assumption is without loss of generality, as the revenue from the first t rounds is then the same, the only difference is that there is already built-up excess demand and supply. However, the structure of built-up excess will play no role in the proof.

I will consider the two cases of weakly monotone sequences separately: (i) a weakly increasing sequence of spreads, and (ii) a weakly decreasing sequence of spreads.

First, I consider the case of a weakly increasing sequence of spreads. It follows from the evolution of demand and supply, that in no round, excess demand and supply is cleared. That is, because a weakly bigger spread corresponds to a larger buy and a smaller sell price. Thus, in every round, only newly incoming traders are matched, that is, the per round revenue is equal to $R^t(\sigma^t) = R^0(\sigma^t)$. Note that for the baseline demand and supply, σ^0 is the optimal spread. Thus, in every round, and hence also for the long-run average, the baseline spread σ^0 weakly improves the revenue and it holds that $R^\infty(\sigma^t) \leq R^\infty(\sigma^0)$. However, it follows from Lemma B.9 that the constant baseline spread is not optimal for long-run average revenue maximization, as it can be strictly improved by the oscillatory spread from the myopic revenue maximization in Theorems 3.1 and 3.2.

Second, I consider the case of a weakly decreasing sequence of spreads. Such a sequence must converge to some spread $\sigma^\infty \geq 0$. Thus, the distance $|\sigma^t - \sigma^{t-1}|$ converges to zero. It was shown in the proof of Theorem 3.1, that the increased per-round revenue $R^t(\sigma^t)$ converges to $R^0(\sigma^t)$: Recall, the reason for this is that a small decrease in the spread only leads to a small decrease in buy and sell prices, and thus only a small fraction of excess demand and supply is cleared. Moreover, as the spreads σ^t converge to σ^∞, it follows from the continuity of revenue in the spread that $R^0(\sigma^t)$ converges to $R^0(\sigma^\infty)$. Hence, there exists a function ϵt that converges to zero, such that $R^t(\sigma^t) \leq R^0(\sigma^\infty) + \epsilon(t)$. Next, we consider two cases separately: (i) $\sigma^\infty \neq \sigma^0$ and (ii) $\sigma^\infty = \sigma^0$.

For (i), that is $\sigma^\infty \neq \sigma^0$, we note that there exists $\delta > 0$, such that $R^0(\sigma^\infty) \leq R^0(\sigma^0) - \delta$, as the spread σ^∞ is not optimal for baseline demand and supply. Hence, for t sufficiently large ($\epsilon(t) \leq \delta$), it holds that $R^t(\sigma^t) \leq R^0(\sigma^\infty) + \epsilon(t) \leq R^0(\sigma^0) - \delta + \epsilon(t) \leq R^0(\sigma^0)$. Thus, for large t, the per-round revenue is weakly smaller than when clearing at the optimal baseline spread σ^0. It follows from Lemma B.9 that from this round onwards, the oscillatory spread from myopic revenue maximization improves the long-run average revenue, which shows that such a sequence is not optimal.

For (ii), that is $\sigma^\infty = \sigma^0$, note that $R^t(\sigma^t) \leq R^0(\sigma^\infty) + \epsilon(t) = R^0(\sigma) + \epsilon(t)$. For every $\epsilon > 0$, take t sufficiently large, such that $\epsilon(t) \leq \epsilon$. Thus, for large t, the per-round revenue improvement from σ^0 to the sequence σ^t is upper bounded by ϵ. Hence, in that large t regime, the improvement on long-run average revenue is also upper bounded by ϵ. Now, it follows from Lemma B.9 that there exists $\delta > 0$, such that the spreads from myopic revenue maximization improve the long-run average revenue compared to the constant spread σ^0 by at least δ. If t is now chosen sufficiently large, such that $\epsilon(t)$ remains smaller than δ, then this shows that in the large t-regime myopic revenue maximization again improves upon the sequence of spreads σ^t. □

B.9 Proof of Theorem A.3

Proof. In a first step, I show that for all σ^T it holds that $\mathcal{A}^T(\sigma^T, \sigma^H) = [\alpha^*(\sigma^T), M]$ for some $\alpha^*(\sigma^T) \in [-M, M]$. For a fixed spread σ^T, consider the map $\alpha \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$. This map is continuous and strictly increasing. To see that the map is continuous, note that $u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)] = u_\alpha(\sigma^T) - \sum_{j\geq 0} u_\alpha(\sigma^j)\mathbb{P}[\sigma^H = \sigma^j]$ holds and the map $\alpha \mapsto u_\alpha(\cdot)$ is continuous. The strict monotonicity follows from the fact that $u_\alpha(\cdot)$ has constant relative risk aversion (CRRA). This implies that as α increases, the function $u_\alpha(\cdot)$ becomes more and more concave. Thus, a trader will prefer the certainty of outcome σ^T over the random nature of the outcome σ^H, proving that the map $\alpha \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$ is indeed strictly increasing.

For fixed σ^T, consider $\alpha^*(\sigma^T)$ such that a trader i is indifferent between the two platforms, that is, $u_\alpha(\sigma^T) = \mathbb{E}[u_\alpha(\sigma^H)]$. This value corresponds to a zero of the map $\alpha \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$. If such a value exists, it is unique, because this map is continuous and strictly increasing. For any $\alpha > \alpha^*(\sigma^T)$, it holds that $u_\alpha(\sigma^T) > \mathbb{E}[u_\alpha(\sigma^H)]$, and thus traders with risk attitude above $\alpha^*(\sigma^T)$ prefer the market with transparent transaction costs. Similarly, for any $\alpha < \alpha^*(\sigma^T)$, it holds that $u_\alpha(\sigma^T) < \mathbb{E}[u_\alpha(\sigma^H)]$, and thus traders with risk attitude below $\alpha^*(\sigma^T)$ prefer the market with hidden transaction costs. Thus, $\mathcal{A}^T(\sigma^T, \sigma^H) = [\alpha^*(\sigma^T), M]$ and $\mathcal{A}^H(\sigma^T, \sigma^H) = [-M, \alpha^*(\sigma^T)]$ holds by construction. If this value does not exist, then for all $\alpha \in [-M, M]$ it either holds that (i) $u_\alpha(\sigma^T) < \mathbb{E}[u_\alpha(\sigma^H)]$ or (ii) $u_\alpha(\sigma^T) > \mathbb{E}[u_\alpha(\sigma^H)]$. In the first case, all traders prefer the market with hidden transaction costs, and it holds that $\mathcal{A}^T(\sigma^T, \sigma^H) = \emptyset$ and $\mathcal{A}^H(\sigma^T, \sigma^H) = [-M, M]$. In that case, I write by abuse of notation that $\alpha^*(\sigma^T) = M$. In the second case, all traders prefer the market with transparent transaction costs, and it holds that $\mathcal{A}^T(\sigma^T, \sigma^H) = [-M, M]$ and $\mathcal{A}^H(\sigma^T, \sigma^H) = \emptyset$. In that case, I write by abuse of notation that $\alpha^*(\sigma^T) = -M$.

Next, I show that $\alpha^*(\sigma^T)$ is continuous and strictly increasing in σ^T. If $\alpha^*(\sigma^T) = M$, then for all $\sigma > \sigma^T$, it holds that $\alpha^*(\sigma) = M$ as well. That is, because $\alpha^*(\sigma^T) = -M$ implies that for all $\alpha \in [-M, M]$ $u_\alpha(\sigma^T) < \mathbb{E}[u_\alpha(\sigma^H)]$. Increasing σ^T to σ only decreases $u_\alpha(\cdot)$. Thus, for all $\alpha \in [-M, M]$ $u_\alpha(\sigma) < \mathbb{E}[u_\alpha(\sigma^H)]$, and hence $\alpha^*(\sigma) = -M$ as well. A similar argument shows that if $\alpha^*(\sigma^T) = -M$, then for all $\sigma < \sigma^T$, it holds that $\alpha^*(\sigma) = -M$ as well. Thus, consider the set of σ^T, where $\alpha^*(\sigma^T) \in (-M, M)$. Recall that $\alpha^*(\sigma^T)$ is the unique zero of the map $\alpha \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$, which is continuous and strictly increasing on $[-M, M]$. Moreover, for fixed $\alpha \in [-M, M]$, the map $\sigma^T \mapsto u_\alpha(\sigma^T) - \mathbb{E}[u_\alpha(\sigma^H)]$ is continuous and strictly decreasing, because for every α, $u_\alpha(\cdot)$ is continuous and strictly decreasing. Then, a standard analytical argument shows that the unique zero $\alpha^*(\sigma^T)$ is continuous and strictly increasing in σ^T.

The revenue from charging a transparent transaction costσ^T is equal to $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H))R^0(\sigma^T)$. Using the above argument, it follows that the first factor is equal to $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H)) = \int_{\alpha^*(\sigma^T)}^{M} f^\alpha d\alpha$, where $f(\alpha)$ is the continuous and strictly positive density function of risk attitudes. Since $\alpha^*(\sigma^T)$ is continuous, it follows that $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H))$ is continuous in σ^T as well. Moreover, for transparent transaction costs, it was shown in Jantschgi et al. (2023) that $R^0(\sigma^{T)}$ is continuous in the spread σ^T as well. Thus, the map $\sigma^T \mapsto \lambda(\mathcal{A}^T(\sigma^T, \sigma^H))R^0(\sigma^T)$ is continuous on $[0, \sigma^0]$.

Note that for any $\sigma \geq \sigma^0$, it follows from the same argument as in Theorem 5.1, that $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H)) = \emptyset$ and thus the revenue is zero. Let $(\sigma^j)_{j\geq 0}$ denote all possible realization of σ^H. Recall from Theorem 3.1 that the this sequence is lower bounded away from 0. For any spread $0 < \sigma < \inf_{j\geq 1} \sigma^H$, it holds that $\mathcal{A}^T(\sigma^T, \sigma^H) = [-M, M]$ and $R^0(\sigma) > 0$. Thus, there exists a spread $\sigma \in (0, \sigma^0)$ such that $\lambda(\mathcal{A}^T(\sigma^T, \sigma^H))R^0(\sigma^T) > 0$. Finally, for $\sigma = 0$, it holds that $R^0(\sigma^T) = 0$. Hence, if an equilibrium exists, it must hold that $\sigma^T \in (0, \sigma^0)$. The existence of a spread $\sigma^T \in (0, \sigma^0)$ with $\sigma^T \in argmax \lambda(\mathcal{A}^T(\sigma^T, \sigma^H))R^0(\sigma^T)$ follows directly from the Extreme Value Theorem. Together with σ^H, this spread σ^T forms an equilibrium. This finishes the proof. $\square$