Toxic Arbitrage and Price Discovery*

Kolja Johannsen†

May 30, 2017

Abstract

Toxic arbitrage opportunities are caused by information arriving in one market leading to short lived price deviations between markets. This paper shows that the direction of such arbitrage opportunities provides valuable insights into price discovery and markets’ information shares. Starting from a new theoretical framework of multi-venue trading, I derive an unbiased measure of information shares based on the frequency of toxic arbitrage opportunities. This measure has several advantages over traditional measures of price discovery, especially when looking at low liquidity environments, and provides a valuable addition in the analysis of price dynamics. I illustrate these advantages with a unique dataset for internationally traded foreign exchange futures.

Keywords: Price Discovery; Arbitrage; Adverse Selection; High Frequency Trading

JEL Classifications: D53, F31, G10

*The author gratefully acknowledges financial support from the ESRC under grant no. 1500669.
†Warwick Business School, University of Warwick, Scarman Road, Coventry, CV4 7AL, UK; kolja.johannsen.14@mail.wbs.ac.uk
1 Introduction

Large parts of the financial system have experienced an increase in market fragmentation and, consequently, there has been an increased interest in market interaction and multi-venue trading. However, most of the existing literature is focusing on highly liquid markets and by doing so is forced to ignore part of the resulting fragmentation. This is partly due to the fact that standard approaches to price discovery reach their limitations when applied to low liquidity environments.

In this paper, I use a new theoretical framework to derive a measure of information share which is robust when applied to markets subject to low liquidity. It therefore provides a useful addition to the analysis of multi-venue trading. The measure builds upon earlier work on toxic arbitrage by Foucault et al. (2016). Toxic arbitrage opportunities are caused by price deviations between markets due to information arrival in one market and asynchronous price adjustment in the other. Such arbitrage opportunities therefore provide useful insights into the information dynamics of markets. This is the first paper to take advantage of the direction of toxic arbitrage opportunities or to use the concept of toxic arbitrage in order to analyse such dynamics.

The contribution to the literature is twofold. Firstly, I introduce a model which highlights the importance of the direction of toxic arbitrage opportunities. This allows me to explore the connection between the informational structure of markets, their spreads and the frequency of toxic arbitrage opportunities. High spreads in one market serve as a protection not only against informed traders in this market but also against arbitrageurs incentivised by informed traders in the other market. Thereby, higher spreads in one market lead to fewer toxic arbitrage opportunities initiated by information arriving in the other market. This connection needs to be taken into account when using the informational content of the frequency of toxic arbitrage opportunities. Similarly, higher transaction costs in either market reduce incentives for informed traders to create toxic arbitrage opportunities. Lastly, the distribution of the impact of information arriving in one market will also need to be taken into account. These connections are important for any future work using toxic arbitrage opportunities. My second contribution and the main contribution of this paper is the introduction of a model based toxic arbitrage information share which corrects for the spreads, transaction costs and informational structures of both markets. This measure is fully theoretically derived and does not share any computational similarities with the standard price discovery shares by Hasbrouck (1995), Gonzalo and Granger (1995) or their extensions. Due to this, the toxic arbitrage information share avoids the criticism brought forward e.g. by Gomber et al. (2016).

To emphasise these contributions, I first compare the toxic arbitrage information share to standard procedures using simulated data. The performance of the toxic arbitrage information share speaks for its theoretical backing and the simulations highlight several advantages of the toxic arbitrage information share especially when looking at low

\[1\] see e.g. Gomber et al. (2016).

\[2\] see e.g. Tse et al. (2006), Brogaard et al. (2014) and Hendershott and Menkveld (2014).
liquidity environments. In a second step, I apply the toxic arbitrage information share to a unique dataset combining US Dollar/Brazilian Real futures traded at the Chicago Mercantile Exchange (CME) and the Bolsa de Valores, Mercadorias e Futuros de São Paulo (BMF). Comparing the results from the toxic arbitrage information share with the results using the Hasbrouck information share and the Gonzalo Granger component share, I find that the three information shares find similar median values. In contrast to the other procedures, the toxic arbitrage information share provides a less volatile and more persistent estimate. In line with the assessment by market participants, the toxic arbitrage information share finds a negligible contribution by CME to price discovery.²

The remainder of this paper is structured as follows. The next section provides a brief overview of the related literature. In Section 3, I derive the theoretical model and develop the toxic arbitrage based information share. Section 4 introduces a procedure of estimating the toxic arbitrage information share. In Section 5, I run several simulations in order to evaluate the performance of the procedure. Sections 6 and 7 provide a description of the dataset used and the empirical application of the toxic arbitrage information share. Section 8 concludes the paper.

2 Literature

This paper forms part of the wide literature of multi-venue trading and connects the concept of toxic arbitrage introduced by Foucault et al. (2016) with the literature on price discovery crucially shaped by Hasbrouck (1995) and Gonzalo and Granger (1995). Research on multi-venue trading has profited greatly by the improved availability of high quality high frequency data and at the same time has become more relevant due to an increased fragmentation across trading venues. Bekaert et al. (2011) find that political risk is one of the main drivers of market fragmentation. The recent debate around capital controls and protectionist policies may therefore lead to a further segmentation. Future research on multi-venue trading should take this into account as a further segmentation is likely to lead to a higher relevance of potentially low liquidity markets.

There is a substantial literature around arbitrage opportunities as a feature of market inefficiency and market frictions. Gromb and Vayanos (2012) provide a broad overview of the theoretical literature. The presence of arbitrageurs can be beneficial or harmful to market participants (see e.g. Copeland and Galai (1983) and Gromb and Vayanos (2002)). Foucault et al. (2016) put this ambiguity at the core of their paper, differentiating between toxic arbitrage and non-toxic arbitrage. While toxic arbitrage is information driven and leads to welfare loss as the arbitrageur takes advantage of the market makers’ latency, non-toxic arbitrage serves as a channel to balance excess liquidity across markets and is therefore welfare enhancing. In this paper, I use this distinction, but rather than focusing on welfare effects, I take advantage of the informational content of toxic arbitrage.

²In discussions with stakeholders of the USD/BRL futures market such as market makers, brokers, traders as well as specialists from BMF and CME, I find that the results using the toxic arbitrage information share are in line with market participants’ expectations that only BMF matters for price discovery.
arbitrage opportunities in order to look at price discovery.

O’Hara (2003) highlights that price discovery is one of the principle purposes of markets. Given that similar or related assets are often traded in multiple venues, a large part of the literature has therefore focused on the analysis of price discovery in order to attribute information shares to different venues. The standard approaches for determining price discovery are the information share going back to Hasbrouck (1995) and the permanent component share by Gonzalo and Granger (1995). Both approaches are based on the Vector Error Correction Model proposed by Engle and Granger (1987). These approaches are the clear benchmark for any analysis of price discovery, but as papers such as Narayan and Smyth (2015) highlight, there are limitations due to their underlying assumptions. Some of these limitations are addressed by adaptations in papers such as Yan and Zivot (2010) and Dias et al. (2016). However, extensions to the standard approaches still rely on complete limit order books or high frequency in trades which are not necessarily given in low liquidity markets. In this paper, I therefore depart from the VECM based price discovery analysis and propose a measure which is based on the frequency of toxic arbitrage opportunities.

Most of the multi-venue literature focuses on equity markets. However, there is also a substantial literature looking at arbitrage and price discovery in foreign exchange markets (see e.g. Tse et al. (2006), Akram et al. (2008), Ranaldo (2009) and Poskitt (2010)). Due to its application to foreign exchange futures, this paper is also contributing to this part of the literature. Ventura and Garcia (2012) find that the futures market in São Paulo accounts for most of the price discovery compared to the spot market. This paper is the first to look at the futures market in Chicago and its relevance for price discovery for the Brazilian Real. The advantage of the toxic arbitrage information share that it does not need high liquidity may be especially beneficial for the analysis of central versus local trading of currencies, a topic of high political interest.

The next section develops the theoretical model which serves several purposes. Firstly, it provides an implication which helps to find a model based estimate of the transaction costs. Further, the model provides a framework for the following analysis and the construction of the toxic arbitrage information share.

3 Model

The following model is inspired by Foucault et al. (2016). Similar to their paper, I am differentiating between the arrival of liquidity traders which potentially leads to non-toxic arbitrage opportunities and informed traders potentially leading to toxic arbitrage opportunities. This model incorporates both types of arbitrage, however the main aim is to explore toxic arbitrage opportunities. While Foucault et al. (2016) focus on the connection between latency and toxic arbitrage, my focus is on the direction and informational impact of toxic arbitrage opportunities. The direction is determined by where the information which leads to the toxic arbitrage opportunity arrives.

For a further description of how the procedures are related, see e.g. Baillie et al. (2002)
There are two markets $B$ and $C$ trading the same asset, three dates ($t \in \{0, 1, 2\}$), two market makers, one in each market, an arbitrageur between markets and both a liquidity and an informed trader in each market. In total, there are seven agents. The central point of this setup is that the arbitrageur is the only market participant able and willing to trade in both markets due to market entry costs. Furthermore, her orders move faster than information on the prices in each market. Market $C$ can be thought of as CME and market $B$ as BMF. Figure 1 provides the time line of the model. The description below is in reverse order.

Figure 1: Time line

\[
\begin{array}{ccc}
\hline
\text{t = 0} & \text{t = 1} & \text{t = 2} \\
\hline
\text{Arrival of informed} & \text{Posting of Prices} & \text{Realisation of the} \\
\text{or liquidity trader} & \text{and Trades} & \text{Final Value } \theta \\
\hline
\end{array}
\]

At $t = 2$ the final value $\theta = \mu + \epsilon$ of the asset is realised, where $\epsilon$ is the variation in the fundamental value of the asset following a symmetric distribution around zero. The expected value of the asset is hence $E(\theta) = \mu + E(\epsilon) = \mu$. Market maker $j \in \{B, C\}$ is specialised in market $j$ and is only active in this market. Due to the physical distance between the market places, the information is not instantaneously available in both markets possibly leading to short lived arbitrage opportunities.

At $t = 1$, market makers simultaneously post ask price $a_j$ and bid price $b_j$ for $j \in \{B, C\}$ such that:

\[
a_j = v_j + \frac{S_j}{2}, \quad \text{and} \quad b_j = v_j - \frac{S_j}{2},
\]

(3.1)

where $v_j$ is market maker $j$’s valuation of the asset and $S_j$ is his bid-ask spread. Quotes are for a fixed quantity, normalised to 1. Market maker $j$’s valuation $v_j$ is determined at time $t = 0$.

There are four possible events in this model:

- News arrival in market $B$
- News arrival in market $C$
- Liquidity trade in market $B$
- Liquidity trade in market $C$

Figure 2 provides an illustration of the probabilities of each event. Each of these events has the potential of causing an arbitrage opportunity. A liquidity trade in market $j$ can lead to a non-toxic arbitrage opportunity initiated in $j$ and news arrival in market $j$ can
lead to a toxic arbitrage opportunity initiated in market \( j \). This paper focuses on the informational content of toxic arbitrage opportunities and their frequency.\(^5\)

Let us consider the cases where an informed trader arrives in one of the markets. This occurs with probability \( \alpha \). The informed trader observes \( \epsilon \), the change in the fundamental value at time \( t = 1 \). Hence, he uses this information to trade before the information becomes public knowledge at time \( t = 2 \). This news arrival via the informed trader can lead to what Foucault et al. (2016) describe as toxic arbitrage.\(^6\) Toxic arbitrage opportunities are not considered welfare improving. The arbitrageur uses his lower latency in order to trade against stale trades by the other market maker. The informed trader arrives at market \( B \) with probability \( \alpha \delta_b \). Let us denote by \( n_B \) the price impact of the information which the informed trader holds. Say the informed trader receives negative information about the asset and we denote this as \( n_B = -n_B^+ \), so that \( n_B^+ > 0 \) is the absolute size of the price impact. It follows, that the true value of the asset which is only observed by the informed trader is given by \( \theta = \mu - n_B^+ \). The informed trader has two ways to taking advantage of his information. He can execute a sell market order and trade directly against market maker \( B \)'s best bid. Additionally, the informed trader also has the option of posting a sell limit order in order to be picked up by the arbitrageur. The informed trader will take advantage of both options as long as it is profitable to do so. In the next step, let us look at the profitability condition in more detail.

When using a market order, the informed trader trades at price \( b_B = \mu - \frac{S_B}{2} \) leading

\(^{5}\)A complete solution of this model is provided in the Appendix.

\(^{6}\)It is a central difference between the model used by the authors and the one used in this paper, that not all information arrival leads to toxic arbitrage opportunities.
to the following profits for the informed trader and market maker \( B \), respectively,

\[
\text{Profit}_{\text{Inf}} = b_{B} - (\mu - n_{B}^{+}) - c = \left( \mu - \frac{S_{B}}{2} \right) - (\mu - n_{B}^{+}) - c \quad (3.2)
\]

\[
= n_{B}^{+} - \frac{S_{B}}{2} - c, \quad (3.3)
\]

\[
\text{Profit}_{\text{MM}} = (\mu - n_{B}^{+}) - b_{B} = \mu - n_{B}^{+} - \left( \mu - \frac{S_{B}}{2} \right) \quad (3.4)
\]

\[
= - \left( n_{B}^{+} - \frac{S_{B}}{2} \right). \quad (3.5)
\]

The informed trader arriving in market \( B \) will only use a market order if this results in a positive profit.\(^7\) Let us denote by \( \bar{n}_{B} \), the price impact of information which leads to a zero profit. It follows that

\[
n_{B} - \frac{S_{B}}{2} - c = 0 \quad (3.6)
\]

\[
\bar{n}_{B} = \frac{S_{B}}{2} + c. \quad (3.7)
\]

For all information where \( n_{B}^{+} < \bar{n}_{B} \), the informed trader arriving in market \( B \) will not execute a market order. In other words, if the price impact in absolute terms \( (n_{B}^{+}) \) does not exceed the threshold, the informed trader will not execute a market order. This shows that both transaction cost \( c \) as well as the spread \( S_{B} \) make it harder for the informed trader to use his private information. Both \( c \) and \( S_{B} \) serve as a protection for market maker \( B \).

Additionally, the informed trader has the option of using a sell limit order. He does so in order to incentivise the arbitrageur to trade against him. This means that the informed trader posts a limit order at the highest price which still makes it profitable for the arbitrageur to trade. This price denoted by \( a_{B}^{T} \) needs to be lower than the best ask \( a_{B} \) by market maker \( B \) so that the informed trader’s limit order is now the best ask in market \( B \).

The price \( a_{B}^{T} \) is hence given by

\[
a_{B}^{T} = a_{B} - \gamma \quad (3.8)
\]

\[
= \mu + \frac{S_{B}}{2} - \gamma, \quad (3.9)
\]

where \( \gamma > 0 \). The informed trader chooses \( \gamma \) in order to incentivise the arbitrageur to trade against her limit order. Consequently, \( \gamma \) is the price improvement the informed trader needs to offer compared to the market maker’s best price in order to make it profitable for the arbitrageur to trade. The arbitrageur will buy at this price if he can make a profit by simultaneously selling the asset in market \( C \). Due to the trades in both

\(^7\) The results for positive information with price impact \( n_{B}^{+} \) are equivalent.
markets, the arbitrageur incurs a total cost of $2c$. The profit of the arbitrageur is then given by

$$\text{Profit}_{\text{Arb}} = a^T_B - b_C - 2c$$

$$= (\mu - \frac{SC}{2}) - (\mu + \frac{SB}{2} - \gamma) - 2c$$

$$= \gamma - \frac{SC + SB}{2} - 2c. \quad (3.12)$$

The informed trader will choose the smallest $\gamma$ which still leads the arbitrageur to engage in arbitrage, i.e.

$$\gamma = \frac{SC + SB}{2} + 2c, \quad (3.13)$$

leading to $\text{Profit}_{\text{Arb}} = 0$. In this case, the informed trader uses the arbitrageur as a channel to trade against market maker $C$. $\gamma$ can also be thought of as the price of the arbitrageurs’ service of providing a channel for indirectly trading against the market maker in the other market. Using a limit order results in the following profits

$$\text{Profit}_{\text{Arb}} = b_C - a^T_B - 2c = \left(\mu - \frac{SC}{2}\right) - \left(\mu + \frac{SB}{2} - \gamma\right) - 2c$$

$$= -\frac{SC + SB}{2} + \gamma - 2c = 0, \quad (3.15)$$

$$\text{Profit}_{\text{InfL}} = a^T_B - (\mu - n_B^+) - c$$

$$= \left(\mu + \frac{SB}{2} - \gamma\right) - (\mu - n_B^+) - c$$

$$= n_B^+ - \frac{SC}{2} - 3c, \quad (3.18)$$

$$\text{Profit}_{\text{MMc}} = (\mu - n_B^+) - b_C = (\mu - n_B^+) - \left(\mu - \frac{SC}{2}\right)$$

$$= -\left(n_B^+ - \frac{SC}{2}\right). \quad (3.20)$$

Given a sufficiently large price impact of the information, each market maker $j$ consequently incurs a loss of $-\left(n_B^+ - \frac{SC}{2}\right)$. The informed trader is trading directly against market maker $B$ and uses the arbitrageur to trade indirectly against market maker $C$. Combining the two sub-profits of the informed trader in Equations (3.3) and (3.18) leads to a total profit of

$$\text{Profit}_{\text{Inf}} = \text{Profit}_{\text{InfM}} + \text{Profit}_{\text{InfL}}$$

$$= \left(n_B^+ - \frac{SB}{2} - c\right) + \left(n_B^+ - \frac{SC}{2} - 3c\right)$$

$$= 2n_B^+ - \frac{SB + SC}{2} - 4c. \quad (3.21)$$
However, following the same logic as above, the informed trader arriving in market B will only use a limit order if it is profitable to do so. The profit for the informed trader of using a limit order is given in Equation (3.18) by

$$\text{Profit}_{Infl} = n_B^* - \frac{S_C}{2} - 3c.$$  

Following the same logic as before, the informed trader will only use a limit order if his profit is positive. Hence, I define the news impact which leads to a zero profit from a limit order as \(\bar{n}_B\). For all information where \(n_B^* < \bar{n}_B\), the informed trader will not post a limit order in market B. It follows that

$$\bar{n}_B^* - \frac{S_C}{2} - 3c = 0 \quad \text{(3.22)}$$

$$\bar{n}_B^* = \frac{S_C}{2} + 3c. \quad \text{(3.23)}$$

Let us now consider the case where an informed trader arrives in market C. The probability of this happening is given by \(\alpha(1 - \delta_b)\). The solutions for this case are symmetric to the case where the informed trader arrives in market B. It follows that

$$\bar{n}_B^* = \frac{S_C}{2} + 3c$$

$$\bar{n}_B^* = \frac{S_C}{2} + 3c. \quad \text{(3.24)}$$

The lowest profitable impact of news for a market order in market \(j\) is given by \(\bar{n}_j\). \(\bar{n}_j^*\) is the lowest profitable impact of news for a limit order. Hence, \(\bar{n}_j\) is the threshold for informed traders arriving in market \(j\) to trade profitably against market maker \(j\). \(\bar{n}_j^*\) is the threshold for informed traders arriving in market \(j\) to use a limit order to be picked up by the arbitrageur. Consequently, \(\bar{n}_j^*\) will determine the number of toxic arbitrage opportunities. Say the price impact of an informed trader’s information arriving in market \(B\) is between \(\bar{n}_B^*\) and \(\bar{n}_B^*\). In this case, the informed trader will only execute a market order but not submit a limit order. The market order will take up liquidity and potentially widen the spread by emptying the best price in market \(B\). This will not lead to an arbitrage opportunity. Arbitrage opportunities can only be caused by the submission of limit orders as a limit order narrows the spread in a market. If the spread becomes sufficiently narrow, this opens an arbitrage opportunity. Market orders can only widen the spread and therefore not lead to arbitrage opportunities.

Comparing \(n_j\) and \(n_j^*\) as displayed in Equations (3.24) and (3.25), I find

$$n_j = \frac{S_j}{2} + c < n_j^* = \frac{S_j}{2} + 3c. \quad \text{(3.26)}$$

\(n_j\) is smaller than \(n_j^*\) for \(j \neq k\) and \(j, k \in \{B, C\}\). This implies that it is easier for market \(j\) to deter an informed trader arriving in market \(k\) from using a limit order than to deter an informed trader arriving in (his) market \(j\) from using a market order. It is easier for a
market maker to defend himself against informed traders in other markets than against informed traders in the same market. The reason for this is that an informed trader using a limit order to incentivise the arbitrageur needs to compensate the arbitrageur for his transaction costs.

3.1 Toxic Arbitrage Information Share

In the next step, I use my model in order to find the true or unbiased information share $IS^*$ using the relative share of toxic arbitrage opportunities. The probability for seeing a toxic arbitrage opportunity initiated in market $B$ and $C$, respectively, is given by

$$\begin{align*}
Prob_{B/C} &= \alpha \delta_b \cdot \text{Prob}(n_B^+ > n_B^*) \\
Prob_{C/B} &= \alpha(1 - \delta_b) \cdot \text{Prob}(n_C^+ > n_C^*) .
\end{align*}$$

(3.27)

(3.28)

This is the probability that the informed trader arrives in the initiating market multiplied by the probability that the price impact of the information is sufficiently large for the informed trader to use a limit order.

It follows that the relative frequency of toxic arbitrage opportunities is given by

$$IS^{B}_{\text{tox}} = \frac{Prob_{B/C}}{Prob_{B/C} + Prob_{C/B}}$$

(3.29)

$$= \frac{1}{1 + \left(\frac{1 - \delta_b}{\delta_b} \cdot \frac{Prob(n_C^+ > n_C^*)}{Prob(n_B^+ > n_B^*)}\right)} .$$

(3.30)

This is the percentage of toxic arbitrage opportunities which is initiated by informed traders arriving in market $B$. The ratio $IS^{B}_{\text{tox}}$ quite clearly differs from the true information share which can be thought of as the cumulative price impact of information arriving in this market relative to the overall information arriving. The latter is given by

$$IS^* = \frac{\delta_b E(n_B^+)}{\delta_b E(n_B^*) + (1 - \delta_b) E(n_C^*)}$$

(3.31)

$$= \frac{1}{1 + \left(\frac{1 - \delta_b}{\delta_b} \cdot \frac{E(n_C^+)}{E(n_B^*)}\right)} .$$

(3.32)

For brevity, let us introduce the following notation

$$\Delta_b = \left(\frac{1 - \delta_b}{\delta_b}\right), \quad \Psi = \left(\frac{\text{Prob}(n_C^+ > n_C^*)}{\text{Prob}(n_B^+ > n_B^*)}\right), \quad \Gamma = \left(\frac{E(n_C^+)}{E(n_B^*)}\right) .$$

(3.33)

$\Delta_b$ indicates how likely it is for news to arrive in market $C$ compared to marked $B$. $\Psi$ is the probability that news arriving in market $C$ lead to a toxic arbitrage opportunity, relative to news arriving in market $B$. A $\Psi$ larger than unity implies that it is more likely...
for an informed trader arriving in market $C$ to cause a toxic arbitrage opportunity than for an informed trader arriving in market $B$. $\Gamma$ is given by the expected impact of news arriving in market $C$ relative to the expected price impact of news arriving in market $B$. A $\Gamma$ larger than unity implies that information from informed traders arriving in market $C$ have a higher expected absolute price impact than information from traders arriving in market $B$. Using this notation, I solve $IS_{\text{tox}}^B$ in Equation (3.30) for the unobservable $\Delta_b$

$$IS_{\text{tox}}^B = \frac{1}{1 + \Delta_b \Psi}$$

$$\Rightarrow \Delta_b = \frac{(1 - IS_{\text{tox}}^B)}{\Psi IS_{\text{tox}}^B}$$

and substitute this in Equation (3.32) for the true information share

$$IS^* = \frac{1}{1 + \Delta_b \Gamma}$$

$$= \frac{1}{1 + \frac{(1 - IS_{\text{tox}}^B)}{\Psi IS_{\text{tox}}^B} \Gamma}.$$  

Equation (3.37) provides a way to correct the difference between the relative toxic arbitrage share and the true information share. In line with the idea that toxic arbitrage opportunities are caused by information arrival, more information arriving in market $B$ leads to a higher $IS_{\text{tox}}^B$. A higher $IS_{\text{tox}}^B$ in turn is associated with a higher information share $IS^*$. Following Equation (3.26), a lower spread in market $B$ leads to a lower threshold $n_C^*$ for information arriving in market $C$ to cause a toxic arbitrage opportunity. A lower $n_C^*$ increases $\text{Prob}(n_C^+ > n_C^*)$ and thereby $\Psi$. Consequently, a lower spread in market $B$ is associated with a higher $IS^*$ given $IS_{\text{tox}}^B$. More intuitively, a higher spread in market $B$ makes it harder for informed traders in market $C$ to take advantage of their information using limit orders. Therefore, the information is less likely to cause an arbitrage opportunity and $IS_{\text{tox}}^B$ will underestimate the true information share $IS^*$.

This specification of $IS^*$ in Equation (3.37) provides us with a toxic arbitrage based information share which only depends on three parameters: $IS_{\text{tox}}^B$, $\Psi$ and $\Gamma$. All of these parameters can be approximated by observable information and, consequently, it is possible to calculate this toxic arbitrage information share. It is worth noting that the identity in Equation (3.37) holds, independently of the distribution of information price impacts. The only underlying assumptions are that the expected average price impact is zero and that the distribution is symmetric.

In the next section, I propose a procedure for estimating the toxic arbitrage information share.

### 4 Estimation Procedure

In order to estimate the toxic arbitrage information share, I need estimates for the three parameters $IS_{\text{tox}}^B$, $\Psi$ and $\Gamma$. The first step of the following procedure regards the
identification of toxic arbitrage opportunities. The second step addresses the estimation of $\Psi$ and $\Gamma$.

4.1 Classification of Toxic Arbitrage Opportunities

An arbitrage opportunity is given when the spread between the best bid in one market and the best ask in the other exceeds the costs of taking advantage of the arbitrage opportunity. Hence, it must be given that

$$b_k - a_j > 2c_k + 2c_j$$

where $k, j \in \{C, B\}$, $k \neq j$, (4.1)

where $c_k$ and $c_j$ are the transaction costs per one asset traded in market $k$ and $j$, respectively. We need to take into account two transactions in each market as the arbitrageur also incurs costs when closing his position in each market after the arbitrage opportunity ceased.

Building on the description of asynchronous price adjustment by Schultz and Shive (2010), Foucault et al. (2016) introduce the concept of toxic arbitrage as arbitrage opportunities due to information arriving in one of the markets. As shown in the model above, such arbitrage opportunities are expected to be initiated by a price movement in the market where the information arrives as the informed trader posts a limit order. Toxic arbitrage opportunities end with a price movement in the market which did not initiate it, as the price in that market adjusts due to arbitrage. Arbitrage due to information implies that the resulting price movement is permanent. Foucault et al. (2016) focus on this fact by classifying toxic arbitrage opportunities as arbitrage opportunities resulting in price movements in both markets in the same direction, i.e.

$$(b_{C;\text{post}} - b_{C;\text{pre}})(a_{B;\text{post}} - a_{B;\text{pre}}) > 0$$

and

$$(a_{C;\text{post}} - a_{C;\text{pre}})(b_{B;\text{post}} - b_{B;\text{pre}}) > 0$$

where $b_{j;\text{post}}$ and $a_{j;\text{post}}$ are the bid and ask prices in market $j \in \{C, B\}$ after the arbitrage opportunity ends and $b_{j;\text{pre}}$ and $a_{j;\text{pre}}$ are the bid and ask prices just before the arbitrage opportunity begins. I am slightly deviating from this definition for two reasons. Firstly, there is a significant amount of arbitrage opportunities where one of these prices does not exist, i.e. where the arbitrage opportunity starts with the first limit order on one side of the limit order book or where the arbitrage opportunity ends with the emptying of one side of the limit order book, especially in CME. In such cases, I am unable to compute the price impact of the arbitrage opportunity. Secondly, when focusing on the price on one side of the limit order book, i.e. bid or ask prices, for a given arbitrage opportunity, the identification of the arbitrage opportunities may be falsified by a change in the spreads in one or both markets during the existence of the arbitrage opportunity. This could be taken care of by taking into account the midprice which, however, brings us back to the problem with partially empty limit order books. In order to avoid these problems, I focus on the chronology of price changes. As mentioned before, a toxic arbitrage opportunity ends due to a price movement in the market which did not cause the deviation in price. For each arbitrage opportunity, I hence record which market’s price movement opens
the arbitrage opportunity and which market’s price movement closes it. In line with
the logic by Foucault et al. (2016), if these two markets are not identical, this arbitrage
opportunity is classified as toxic. This means that the price deviation was caused by
information arriving in one market and closed by the asynchronous price adjustment in
the other market. If these two markets are identical, I classify this arbitrage opportunity
as non-toxic, i.e. as liquidity driven. For the remainder of this paper, I am using the
chronology of price movements for the identification of toxic arbitrage opportunities.

4.2 Estimation of the Toxic Arbitrage Information Share

IS\textsuperscript{B}_{tox} is given by the ratio of toxic arbitrage opportunities

\[ IS\textsuperscript{B}_{tox} = \frac{\#_{\text{Toxic BMF/CME}}}{\#_{\text{Toxic BMF/CME}} + \#_{\text{Toxic CME/BMF}}}. \]  

In order to get an approximation for the relative probability for seeing a toxic arbitrage
opportunity in market C relative to market B (Ψ) and the average price impact of
information arriving in market C relative to market B (Γ), I need to make an assumption
about the distribution of the price impacts in each market. Similar to standard
econometric techniques, I assume a normal distribution of price impacts in each market

\[ n_B \sim N(0, \sigma^2_B) \quad \text{and} \quad n_C \sim N(0, \sigma^2_C), \]  

(4.5)

where \(\sigma_j\) is the standard deviation the price impact of information arriving in market
j. The mean price impact is zero. Further, I assume the informational structure not to
change in expectation. In other words, the market makers expectation of \(\sigma_B\) and \(\sigma_C\) do
not change over time. In order to calculate \(\sigma_B\), I am using the observed price impacts
of information arriving in market B given by the size of toxic arbitrage opportunities
initiated in market B. Let us define \(n^+_j\) as the absolute price impact of information
arriving in market j given by

\[ n^+_j = |p_{B;\text{post}} - p_{B;\text{pre}}|, \]  

(4.6)

where \(p_{B;\text{post}}\) is the midprice in market B just after the arbitrage opportunity closed
and \(p_{B;\text{pre}}\) is the midprice in market B just before the arbitrage opportunity opened.
For each toxic arbitrage opportunity, I can estimate this price impact if both midprices
exist. I am using the midprices in market B due to its higher liquidity.

Given the assumption of normality, the distribution of positive observed price impacts
is essentially a truncated normal distribution. The threshold of the truncation is given
by the spread \(S_C\) and the transaction cost for all relevant transactions. There are
five relevant transactions in every case.\(^8\) The expected value of the truncated normal
distribution subject to this threshold is illustrated in Figure 3. Formally, the expected

\(^8\) As in the calculation of the arbitrage opportunities, I need to take into account all transaction costs
which the arbitrageur is facing. The arbitrageur needs to sell the asset after correction of the arbitrage
opportunity and hence needs to take into account the costs for these trades as well. Additionally, it is
necessary to include the costs for the submission of a limit order by the informed trader. This makes
five transactions in total.
value of the truncated normal distribution is given by

$$E(n_B^+|n_B > n_B^*) = \sigma_B \phi \left( \frac{n_B^*}{\sigma_B} \right) \left( 1 - \Phi \left( \frac{n_B^*}{\sigma_B} \right) \right),$$  

(4.7)

where $\phi(\cdot)$ is the probability density function of the standard normal distribution and $\Phi(\cdot)$ is its cumulative distribution function

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} x^2 \right), \quad \Phi(x) = \frac{1}{2} \left( 1 + \text{erf}(x/\sqrt{2}) \right).$$  

(4.8)

The average value of the observed positive price impacts should be asymptotically equal to the expected value of the truncated normal distribution subject to this threshold. By setting $E(n_B^+|n_B > n_B^*)$ equal to the mean of all positive price impacts of $B/C$ arbitrage opportunities, I am hence able to derive $\sigma_B$.\textsuperscript{9} In order to do so, I numerically solve for $\sigma_B$ which solves the equation

$$\text{mean}(n_B^+) = \frac{\sigma_B \phi \left( \frac{n_B^*}{\sigma_B} \right)}{\left( 1 - \Phi \left( \frac{n_B^*}{\sigma_B} \right) \right)}.$$  

(4.9)

This provides us with an estimate of $\sigma_B$, the standard deviation of price impacts of information arriving in market $B$. Given the assumption of a normal distribution with

\textsuperscript{9}I am using the absolute price impact of all $B/C$ arbitrage opportunities here. Given the assumption that the underlying distribution of price impacts is symmetric and around zero, this does not influence the results but provides us with more observed price impacts.
zero mean, the distribution of price impacts is fully identified. I can now use this distribution to estimate the denominators of $\Psi$ and $\Gamma$ as shown in Equation (3.33). The denominator of $\Psi$ is the probability for seeing a $B/C$ arbitrage opportunity conditional on that an informed trader arrives in market $B$. This is equal to one minus the cumulative distribution function of a normal distribution given $\sigma_B$

\[
 Prob\left(n_B^+ > \bar{n}_B^*\right) = 1 - Prob\left(n_B^- < \bar{n}_B^*\right) \\
= 1 - \Phi\left(\frac{n_B^*}{\sigma_B}\right). \tag{4.10}
\]

The denominator of $\Gamma$ is the expected positive price impact of information arriving in market $B$, which is equivalent to the expected value of a normal distribution truncated at zero

\[
 E\left(n_B^+\right) = E\left(n_B|n_B > 0\right) \tag{4.12}
= \frac{\sigma_B\phi\left(0\right)}{\left(1 - \Phi\left(0\right)\right)} \tag{4.13}
= \frac{2\sigma_B}{\sqrt{2\pi}}. \tag{4.14}
\]

This is illustrated in Figure 4.

Figure 4: Normal distribution truncated at zero

Following the same steps as above, I can also derive the numerators of $\Psi$ and $\Gamma$ by determining the standard deviation of the distribution of price impacts of information arriving in market $C$. This leads to

\[
 Prob\left(n_C^+ > \bar{n}_C^*\right) = 1 - \Phi\left(\frac{n_C^*}{\sigma_C}\right) \quad \text{and} \quad E\left(n_C^+\right) = \frac{2\sigma_C}{\sqrt{2\pi}}. \tag{4.15}
\]
From here, I am able to identify $\Psi$ and $\Gamma$ as

$$
\Psi = \frac{1 - \Phi \left( \frac{\sigma_C}{\sigma_B} \right)}{1 - \Phi \left( \frac{\sigma_B}{\sigma_B} \right)} \quad \text{and} \quad \Gamma = \frac{\sigma_C}{\sigma_B}
$$

(4.16)

When combining this with $IS_{\text{tox}}^B$ as in Equation (4.4), I am able to fully identify $IS^*$ using Equation (3.37). I denote the resulting measure as toxic arbitrage information share (TAIS)

$$
TAIS = IS^*
$$

(4.17)

$$
= \frac{1}{1 + \left( \frac{1 - IS_{\text{tox}}^B}{IS_{\text{tox}}^B} \right) \Psi^{-1} \Gamma}
$$

(4.18)

$$
= \left( 1 + \frac{1 - IS_{\text{tox}}^B}{IS_{\text{tox}}^B} \right) \Psi^{-1} \Gamma
$$

(4.19)

$$
= \left( 1 + \frac{1 - IS_{\text{tox}}^B}{IS_{\text{tox}}^B} \right) \frac{1 - \Phi \left( \frac{\sigma_B}{\sigma_B} \right)}{1 - \Phi \left( \frac{\sigma_B}{\sigma_B} \right)}
$$

(4.20)

where the thresholds are given by $\bar{n}_B^* = S_C/2 + 5c$ and $\bar{n}_C^* = S_B/2 + 5c$.10

In order to calculate the toxic arbitrage information share, I need to take into account that the spreads are not constant over time. This makes it necessary to be more explicit about the nature of $S_j$. I denote by $S_{j,\text{pre}}$ the spread in market $j$ just before an arbitrage opportunity is initiated in the other market. Hence, $S_{B,\text{pre}}$ is the spread in market $B$ just before information arrives in market $C$. This spread in market $j$ together with the transaction cost determine the threshold whether information arrival in the other market leads to a toxic arbitrage opportunity. In order to determine the underlying distribution of price impacts from the observed toxic arbitrage opportunities, the different thresholds need to be taken into account. For each $S_{i,B} \in S_{B,\text{pre}}$, I am numerically approximating11 the value of $\sigma_{i,C}$, the standard deviation of the price impacts from information arriving in market $C$, following

$$
\text{mean}(n_{C}^+|S_{B,\text{pre}} = S_{i,B}) = \frac{\sigma_{i,C} \Phi \left( \frac{\bar{n}_C}{\sigma_C} \right)}{1 - \Phi \left( \frac{\bar{n}_C}{\sigma_C} \right)},
$$

(4.21)

where $\bar{n}^*_C$ is the threshold for information arriving in market $C$ to cause a toxic arbitrage opportunity given a spread of $S_{i,B}$ in market $B$. For each $S_{i,B}$ there can be several estimates of $\sigma_C$ if there are multiple solutions to the equation. All of the resulting

---

10 As mentioned before, the thresholds are the sum of the halfspread and the total transaction cost involved in the creation of an arbitrage opportunity.

11 I am following the procedure by Soetaert and Herman (2009).
estimates $\sigma_{i,C}$ are approximations of the same $\sigma_C$ of the underlying distribution. In order to take all of these into account, I calculate the weighted average of the standard deviations

$$\sigma_C = \sum \frac{w_i}{\sum w_i} \sigma_{i,C}. \quad (4.22)$$

The weights $w_i$ are given by the number of $C/B$ toxic arbitrage opportunities at each $S_{i,B}$. Following the same procedure, I am estimating $\sigma_B$. Using these estimates for the standard deviations, it is possible to estimate $\Gamma$ as

$$\Gamma = \frac{\sigma_C}{\sigma_B}. \quad (4.23)$$

$\Gamma$ is constant as long as the informational structure of the markets ($\sigma_C$ and $\sigma_B$) does not change. Given the standard deviations of the price impact of information, the next step is to calculate $\Psi$ given by

$$\Psi = \left( \frac{\text{Prob} \left( n_{i,C}^+ > \bar{n}_{i,C}^* \right) }{\text{Prob} \left( n_{i,B}^+ > \bar{n}_{i,B}^* \right) } \right). \quad (4.24)$$

$\Psi$ depends on the thresholds $\bar{n}_{i,C}^*$ and $\bar{n}_{i,B}^*$ and therefore changes with the spreads in both markets. Let us denote by $\Psi_t$ the value of $\Psi$ at a specific point in time $t$. The spreads at this time are given by $\sigma_{t,C}$ and $\sigma_{t,B}$. For each $t$, I can calculate $\Psi_t$ as given by

$$\Psi_t = \frac{1 - \Phi \left( \frac{n_{i,C}^+}{\sigma_{t,C}} \right) }{1 - \Phi \left( \frac{n_{i,B}^+}{\sigma_{t,B}} \right) }, \quad (4.25)$$

where $n_{i,C}^*$ and $n_{i,B}^*$ are the thresholds given spreads $\sigma_{t,C}$ and $\sigma_{t,B}$, respectively.\(^{12}\) I then use the average of the $\Psi_t$ as an estimate for the $\Psi$

$$\Psi = \frac{1}{T} \sum_{t=1}^{T} \Psi_t. \quad (4.26)$$

In the next section, I am using the procedure outlined above in order to compare the performance of standard procedures with the toxic arbitrage information share.

### 4.3 Source of Adverse Selection

In this model, there are two sources of adverse selection for each market maker. A market maker can be taken advantage of by an informed trader in the same market. Additionally, a market maker’s stale quote can be traded against by an arbitrageur.\(^{12}\) Alternatively, one can take the average of $\text{Prob} \left( n_{i,C}^+ > n_{i,C}^* \right)$ and $\text{Prob} \left( n_{i,B}^+ > n_{i,B}^* \right)$ separately, before dividing one by the other. The results differ due to Jensen’s inequality, however, the difference is negligible.
The frequency with which informed traders arrive in market \( B \) is given by \( \alpha \delta_B \). The model does not allow us to solve for \( \alpha \), however from Equation (3.32) we know that

\[
IS^* = \frac{\delta_b E(n_B^+)}{\delta_b E(n_B^+) + (1 - \delta_b) E(n_C^+)}
\]  
(4.27)

\[
\delta_b = \frac{IS^* E(n_C^+)}{E(n_B^+) + IS^*(E(n_C^+) - E(n_B^+))}.
\]  
(4.28)

Given that \( E(n_B^+) = 2 \frac{\sigma_B}{\sqrt{2\pi}} \) and \( E(n_C^+) = 2 \frac{\sigma_C}{\sqrt{2\pi}} \), this leads to

\[
\delta_b = \frac{IS^* E(n_C^+)}{E(n_B^+) + IS^*(E(n_C^+) - E(n_B^+))}
\]  
(4.29)

\[
\delta_b = \frac{IS^* 2 \frac{\sigma_C}{\sqrt{2\pi}}}{E(n_B^+) + IS^* (2 \frac{\sigma_C}{\sqrt{2\pi}} - 2 \frac{\sigma_B}{\sqrt{2\pi}})}
\]  
(4.30)

\[
\delta_b = \frac{IS^* \sigma_C}{\sigma_B + IS^*(\sigma_C - \sigma_B)}.
\]  
(4.31)

Furthermore, the probability with which market maker \( B \) is subject to adverse selection via the arbitrageur is given by

\[
Prob_{C/B} = \alpha (1 - \delta_b) \cdot Prob(n_C^+ > n_C^+)
\]  
(4.32)

while the probability that market maker \( B \) is taken advantage of by an informed trader arriving in market \( B \) is given by

\[
Prob_B = \alpha \delta_b \cdot Prob(n_B^+ > n_B^+).
\]  
(4.33)

The ratio of the two indicates how often market maker \( B \) is subject to adverse selection due to latency:

\[
Adv_b = \frac{(1 - \delta_b) \cdot Prob(n_C^+ > n_C^+)}{\delta_b \cdot Prob(n_B^+ > n_B^+)}
\]  
(4.34)

\[
Adv_b = \frac{(1 - \delta_b) \cdot 1 - \Phi \left( \frac{n_C^+}{\sigma_C} \right)}{\delta_b \cdot 1 - \Phi \left( \frac{n_B^+}{\sigma_B} \right)}.
\]  
(4.35)

Suppose we find an \( Adv_b = 2 \).\textsuperscript{13} This implies that market maker \( B \) is twice as often subject to adverse selection due to latency than due to informed traders in market \( B \). Due to symmetry, the adverse selection ratio for market maker \( C \) is given by

\[
Adv_c = \frac{\delta_b \cdot Prob(n_B^+ > n_B^+)}{(1 - \delta_b) \cdot Prob(n_C^+ > n_C^+)}
\]  
(4.36)

\[
Adv_c = \frac{\delta_b \cdot 1 - \Phi \left( \frac{n_B^+}{\sigma_B} \right)}{(1 - \delta_b) \cdot 1 - \Phi \left( \frac{n_C^+}{\sigma_C} \right)}.
\]  
(4.37)

\textsuperscript{13}Adv_b is only defined if the denominator is not zero.
Both ratios depend not only on the frequency of information arriving in each market determined by $\delta_b$ but also on the price impact distributions of the information entering both markets determined by $\sigma_C$ and $\sigma_B$ as well as the spreads and transaction costs which shape $\bar{n}^*_C$ and $\bar{n}_B$ as shown in Equations (3.26)

$$n_j = \frac{S_j}{2} + c < n^*_k = \frac{S_j}{2} + 3c$$

where $j, k \in \{B, C\}$ and $j \neq k$. A higher proportion of information arriving in market $B$ implies a lower adverse selection ration in market $B$ and a higher $\text{Adv}_b$. In this case, market maker $B$ faces more direct adverse selection and less adverse selection due to latency. The reverse is true for market maker $C$. If the price impact of information arriving in market $B$ is higher (a higher $\sigma_B$), $\text{Adv}_b$ will be lower and $\text{Adv}_c$ higher. As the price impact has a higher standard deviation, the average price impact is higher and it is more likely for information arriving in market $B$ to be profitable to the informed trader. As the Gauss error function $\text{erf}(x)$ has a positive first derivative, these properties are also true for the cumulative distribution function $\Phi(x)$. If we assume that the information in both markets has the same expected price impact, it follows that $\frac{\frac{S}{2} + 3c}{\sigma_B} > \frac{\frac{S}{2} + c}{\sigma_B}$ and as $\frac{\partial \Phi(x)}{\partial x} > 0$

$$\frac{1 - \Phi \left( \frac{\frac{S}{2} + 3c}{\sigma_B} \right)}{1 - \Phi \left( \frac{\frac{S}{2} + c}{\sigma_B} \right)} < 1.$$ 

(4.39)

This implies that if $\sigma_C = \sigma_B$, the adverse selection ratio is lower than the ratio of the probabilities of information arriving: $\text{Adv}_b < \frac{(1 - \delta_b)}{\delta_b}$. This once more highlights that a market maker is better protected from information in the other market than from an informed trader arriving in the same market. The difference between the exposure to the two informed traders is due to the transaction costs. Higher transaction costs make this effect stronger.

A change in the spreads has two opposing effects. A wider spread in market $B$ makes it less likely that the informed trader can profit from trading with market maker $B$. At the same time, it is less profitable for an informed trader arriving in market $C$ to use a limit order to incentivise an arbitrageur to trade in both markets. The first effect will lead to a higher $\text{Adv}_b$ while the second will lead to a lower $\text{Adv}_b$. $\text{Adv}_c$ is unaffected by a change in the spread in market $B$. Which effect dominates depends on the combination of parameters.

5 Simulations

The toxic arbitrage information share provides a model based measure of information arriving in each market. It is reasonable to expect this to be similar in size as other measures of price discovery. The standard procedures for price discovery are given by the
information share \((IS)\) introduced by Hasbrouck (1995) and the permanent component share \((CS)\) by Gonzalo and Granger (1995). For a detailed description and implementation of these procedures see e.g. Tse et al. (2006), Chen and Gau (2010) and Yan and Zivot (2010).

As highlighted by Baillie et al. (2002), the computation of the Gonzalo Granger and Hasbrouck procedures are closely related as both are based on Vector Autoregressive Regressions (VAR). Yan and Zivot (2010) and Baillie et al. (2002) highlight the importance of choosing the frequency when using these procedures. A lower frequency is expected to lead to wider bands of the Hasbrouck information share, while higher frequency data may be affected more by microstructure noise. It is one of the central advantages of the toxic arbitrage information share that the frequency of the data does not matter. It is worth noting that the toxic arbitrage information share purely focuses on the arrival of information in each market. The information shares introduced by Gonzalo and Granger (1995) and Hasbrouck (1995) as well as their numerous extensions have a less narrow definition. The Gonzalo Granger permanent component share measures the relative influence of the underlying equilibrium price on each market. It is not straightforward, however, how a market would be able to adjust a price towards the long-run equilibrium without some sort of information arrival. When looking at illiquid markets, there are further problems with data availability. Of course, all procedures are unable to take into account periods where one of the limit order books is empty. However, the toxic arbitrage information share is more flexible in this regard as it can handle limit order books which are partially empty. Narayan and Smyth (2015) provide an overview of the assumptions and shortcomings of the standard procedures resulting from a simple implementation of the VAR approach. For example, the VAR system depends on the correct identification of the lag structure and that the structure does not change over the estimation period. The VAR is also more prone to errors due to missing observations during the day which interrupt the time series. In these regards, the toxic arbitrage information share proposed here is more flexible and robust.

In the remainder of this section, I am running several simulations in order to compare the performance of the toxic arbitrage information share with the Hasbrouck information share as well as the Gonzalo Granger permanent component share.

5.1 Simulation I

The efficient price \(m_t\) is given by a modified random walk

\[
m_t = m_{t-1} + \hat{u}_t
\]

\[
\hat{u}_t = 1 \mu u_t,
\]
where \( u_t \) are price innovations following a discretised normal distribution \( u_t \sim \mathcal{N}(0, \sigma_u^2) \).\(^{14}\) In order to take into account periods with no price innovation, I introduce \( 1_p \) which is a binomially distributed dummy which is one with probability \( p \) and zero otherwise. If \( p = 0.4 \), this implies that there is a change in the efficient price 40% of the time and 60% of the time there is not. Similarly, the efficient price in market \( j \in \{B,C\} \) is given by

\[
m_{j,t} = m_{t-1} + \hat{u}_{j,t}.
\]

(5.3)

The set of price innovations in market \( B \) (\( \hat{u}_{B,t} \)) is a subset of the overall price innovations \( \hat{u}_t \). This subset affects market \( B \) immediately, while the remaining set of \( \hat{u}_t \) only affects the efficient price in market \( B \) with a lag via \( m_{t-1} \). Formally, this can be described as

\[
\hat{u}_{B,t} = 1_{p_B} \hat{u}_t,
\]

(5.4)

where \( 1_{p_B} \) is a binomially distributed dummy which is 1 with probability \( p_B \). Whenever an innovation is not immediately affecting the efficient price in market \( B \), it is immediately affecting the efficient price in market \( C \). Hence, \( \hat{u}_{C,t} \) is given by

\[
\hat{u}_{C,t} = (1 - 1_{p_B}) \hat{u}_t.
\]

(5.5)

This implies that \( p_B \) is the information share of market \( B \). Finally, the observed best prices in market \( j \) are given by

\[
a_{j,t} = m_{j,t} + \frac{S_{j,t}}{2} \quad \text{(5.6)}
\]

\[
b_{j,t} = m_{j,t} - \frac{S_{j,t}}{2} \quad \text{(5.7)}
\]

The following simulations are based on 10,000 observations with \( p = 0.5 \) and \( \sigma_u = 1 \). In order to evaluate the precision of the different price discovery measures, I estimate them for varying \( p_B \). The Hasbrouck information share and the permanent component share are based on the observed midprice

\[
\text{obs}_j = b_{j,t} + \frac{a_{j,t} - b_{j,t}}{2}.
\]

(5.8)

The toxic arbitrage information share is estimated by applying the procedure outlined in Section 4 using the observed best prices \( a_{j,t} \) and \( b_{j,t} \).

Figure 5a and 5b illustrate the results for the toxic arbitrage information share (\( TAIS \)), the Gonzalo Granger permanent component share (\( CS \)), and the Hasbrouck information share (\( IS \)) simulating the above price series 20 times for different values of

---

14 In order to take into account that prices can only change in multiples of the ticksize, I am discretising all random variables in the simulation by rounding them to 2 digits. This is equivalent to saying that the ticksize is 0.01. The discretisation is important for this exercise in order to ensure that there is a limited number of spread sizes, which in turn is necessary for the calculation of the toxic arbitrage information share.
Figure 5: Simulation I

The simulations in Figure 5a are based on $S_{B,t} = S_{C,t} = 2$ and the simulations in Figure 5b use $S_{B,t} = 1$ and $S_{C,t} = 2$. The x-axis is given by the true underlying $p_B$, while the y-axis is given by the estimated information shares. The black 45 degree line hence marks an unbiased estimate. The red line shows the median toxic arbitrage information share resulting from the simulations. The red shaded error bands show the area engulfing 90% of the estimates. The results for the permanent component share are given in blue. The green shaded area illustrates the area between the median upper bound and the median lower bound of the Hasbrouck information share. The green line illustrates the midpoint of this band.

Figure 5a shows that the bounds of the Hasbrouck information share (green) engulf the 45 degree line for intermediate values of $p_B$ implying that it correctly measures the information share for such values. For $p_B < 0.2$ or $p_B > 0.8$, the true information share lies outside the bounds almost every time. All three measures correctly identify the information share around $p_B = 0.5$. Overall, the toxic arbitrage information share performs better than the permanent component share but underperforms compared to Hasbrouck. The permanent component share appears unable to differentiate information shares between 20% and 80%. Figure 5b broadly confirms these results. The reduction in $S_B$ and the resulting difference in spreads appears to mostly affect the accuracy of the toxic arbitrage information share. It’s performance significantly improves for high $p_B$ but becomes worse for low and intermediate $p_B$. The Hasbrouck information share is performing slightly worse than the toxic arbitrage information share for high levels of $p_B$. 

$p_B$. The simulations in Figure 5a are based on $S_{B,t} = S_{C,t} = 2$ and the simulations in Figure 5b use $S_{B,t} = 1$ and $S_{C,t} = 2$. The x-axis is given by the true underlying $p_B$, while the y-axis is given by the estimated information shares. The black 45 degree line hence marks an unbiased estimate. The red line shows the median toxic arbitrage information share resulting from the simulations. The red shaded error bands show the area engulfing 90% of the estimates. The results for the permanent component share are given in blue. The green shaded area illustrates the area between the median upper bound and the median lower bound of the Hasbrouck information share. The green line illustrates the midpoint of this band.

Figure 5a shows that the bounds of the Hasbrouck information share (green) engulf the 45 degree line for intermediate values of $p_B$ implying that it correctly measures the information share for such values. For $p_B < 0.2$ or $p_B > 0.8$, the true information share lies outside the bounds almost every time. All three measures correctly identify the information share around $p_B = 0.5$. Overall, the toxic arbitrage information share performs better than the permanent component share but underperforms compared to Hasbrouck. The permanent component share appears unable to differentiate information shares between 20% and 80%. Figure 5b broadly confirms these results. The reduction in $S_B$ and the resulting difference in spreads appears to mostly affect the accuracy of the toxic arbitrage information share. It’s performance significantly improves for high $p_B$ but becomes worse for low and intermediate $p_B$. The Hasbrouck information share is performing slightly worse than the toxic arbitrage information share for high levels of $p_B$. 

$p_B$. The simulations in Figure 5a are based on $S_{B,t} = S_{C,t} = 2$ and the simulations in Figure 5b use $S_{B,t} = 1$ and $S_{C,t} = 2$. The x-axis is given by the true underlying $p_B$, while the y-axis is given by the estimated information shares. The black 45 degree line hence marks an unbiased estimate. The red line shows the median toxic arbitrage information share resulting from the simulations. The red shaded error bands show the area engulfing 90% of the estimates. The results for the permanent component share are given in blue. The green shaded area illustrates the area between the median upper bound and the median lower bound of the Hasbrouck information share. The green line illustrates the midpoint of this band.

Figure 5a shows that the bounds of the Hasbrouck information share (green) engulf the 45 degree line for intermediate values of $p_B$ implying that it correctly measures the information share for such values. For $p_B < 0.2$ or $p_B > 0.8$, the true information share lies outside the bounds almost every time. All three measures correctly identify the information share around $p_B = 0.5$. Overall, the toxic arbitrage information share performs better than the permanent component share but underperforms compared to Hasbrouck. The permanent component share appears unable to differentiate information shares between 20% and 80%. Figure 5b broadly confirms these results. The reduction in $S_B$ and the resulting difference in spreads appears to mostly affect the accuracy of the toxic arbitrage information share. It’s performance significantly improves for high $p_B$ but becomes worse for low and intermediate $p_B$. The Hasbrouck information share is performing slightly worse than the toxic arbitrage information share for high levels of $p_B$. 

$p_B$. The simulations in Figure 5a are based on $S_{B,t} = S_{C,t} = 2$ and the simulations in Figure 5b use $S_{B,t} = 1$ and $S_{C,t} = 2$. The x-axis is given by the true underlying $p_B$, while the y-axis is given by the estimated information shares. The black 45 degree line hence marks an unbiased estimate. The red line shows the median toxic arbitrage information share resulting from the simulations. The red shaded error bands show the area engulfing 90% of the estimates. The results for the permanent component share are given in blue. The green shaded area illustrates the area between the median upper bound and the median lower bound of the Hasbrouck information share. The green line illustrates the midpoint of this band.
5.2 Simulation II

The second simulation introduces noise around the bid and ask of each market. In the second price generating process, the efficient price $m_t$ is again following the modified random walk

$$m_t = m_{t-1} + \hat{u}_t$$

$$\hat{u}_t = 1_p u_t$$

and the efficient price in market $j$ is given by

$$m_{j,t} = m_{t-1} + \hat{u}_{j,t}$$

$$\hat{u}_{B,t} = 1_p \hat{u}_t$$

$$\hat{u}_{C,t} = (1 - 1_p) \hat{u}_t.$$ 

The difference to simulation I is in the processes of the best bid and ask in market $j$

$$a_{j,t} = m_{j,t} + \frac{S_j}{2} + 1_p \epsilon_{a_j,t}$$

$$b_{j,t} = m_{j,t} - \frac{S_j}{2} + 1_p \epsilon_{b_j,t}.$$ 

where $1_p$ for $j \in \{B, C\}$ are binomially distributed dummies which is one with probability $p_{\epsilon,j}$. These dummies determine the frequency with which noise in the observed prices appears. $\epsilon_{a_j,t}$ and $\epsilon_{b_j,t}$ are uncorrelated and follow discretised normal distributions $\epsilon_{a_j,t} \sim N(0, \sigma^2_{\epsilon,j})$ and $\epsilon_{b_j,t} \sim N(0, \sigma^2_{\epsilon,j})$. As before, the discretisation ensures that the prices move within the grid of ticks.

Figure 6: Simulation II
The following simulations are based on 10,000 observations with \( p = 0.5, \sigma_u = 1, \sigma_u = 1, S_{B,t} = 1, \) and \( S_{C,t} = 2. \) The standard deviations of the noise are given by \( \sigma_{e,B} = \sigma_{e,C} = 0.2. \) The frequency of the noise is set to \( p_{e,B} = p_{e,C} = 0.2. \) Consequently, at each point in time there is a 20% probability for each price to be affected by noise. Figure 6 illustrates the results. The error bands around the toxic arbitrage information share and the permanent component share have widened as a result of the noise in bid and ask prices. The Hasbrouck information share appears to be performing better for an information share around 95%. Apart from this, the previous results are not affected.

### 5.3 Simulation III

The third price generating process follows the process in simulation I. However, it introduces missing observations in the bid and ask prices of market \( C. \) Most of the price generating process of the first simulation is hence unchanged

\[
m_t = m_{t-1} + \hat{u}_t \\
\hat{u}_t = \mathbb{1}_p u_t \\
m_{j,t} = m_{t-1} + \hat{u}_{j,t} \\
\hat{u}_{B,t} = \mathbb{1}_p \hat{u}_t \\
\hat{u}_{C,t} = (1 - \mathbb{1}_p \hat{u}_t).
\]

The difference lies in how the best bid and ask react to the efficient price in market \( j \)

\[
a_{j,t} = \begin{cases} 
  m_{j,t} + \frac{S_{j,t}}{2} & \text{with probability } 1 - p_{a,j} \\
  \text{NA} & \text{with probability } p_{a,j}
\end{cases} 
\]

\[
b_{j,t} = \begin{cases} 
  m_{j,t} - \frac{S_{j,t}}{2} & \text{with probability } 1 - p_{b,j} \\
  \text{NA} & \text{with probability } p_{b,j}
\end{cases}
\]

where \( p_{b,j} \) and \( p_{a,j} \) is the probability for a missing value in the best bid and ask of market \( j, \) respectively.

The following simulations are based on 10,000 observations with \( p = 0.5, \sigma_u = 1, \sigma_u = 1, S_{B,t} = 1, S_{C,t} = 2, p_{a,B} = p_{b,B} = 0, \) and \( p_{a,C} = p_{b,C} = 0.1. \) Hence, ten percent of the bid and ask prices, respectively, of market \( C \) are missing. I find that the results of the Hasbrouck information share worsen both in terms of the accuracy as well as the precision of the measurement. The permanent component share performs better than the midpoint of the Hasbrouck information share and provides a fairly small error band. In fact, \( CS \) appears to be performing better in this simulation than in any of the previous ones. The toxic arbitrage information share shows small error bands and provides the best measurement for \( p_B > 0.65 \) and worse otherwise.

The precision of the toxic arbitrage information share crucially depends on the number of toxic arbitrage opportunities in the sample. Figures 8a) and 8b) illustrate this for a true information share of 95% and 80%, respectively. The simulations are based on the same specification as the previous simulation. The x-axis is given by the number of toxic arbitrage opportunities. The red line depicts the median \( TAIS \) for a given number of
toxic arbitrage opportunities. 90% of the estimated TAIS lie within the light red area. Ten toxic arbitrage opportunities provide a similar median accuracy as higher numbers, however, the error band stretches over 30 and 40 percentage points, respectively. With 50 toxic arbitrage opportunities, the band is reduced to a width of 10 percentage points.

In summary, none of the procedures appears clearly superior to the other. The result is similar to the results by Hasbrouck (2002) using simulations to compare IS and CS.
This suggests that the three measures provide good complements in analysing the price discovery dynamics. The toxic arbitrage information share introduced in this paper is performing especially well for situations where most of the price discovery takes place in the more liquid market. For such cases, it performs better than the alternatives. Furthermore, I find that an increase in the number of toxic arbitrage opportunities quickly improves the performance of the measure. It appears reasonable to focus on samples with at least ten toxic arbitrage opportunities.

In the following sections, I apply this new measure of information shares to a unique high frequency dataset of foreign exchange futures. Section 6 provides a description of the data while Section 7 shows the empirical application.

6 Description of Market Setup and Data

For the application of the toxic arbitrage information share, I am using a unique dataset combining high frequency data from Chicago Mercantile Exchange (CME) and Bolsa de Valores, Mercadorias e Futuros de São Paulo (BMF). BMF is the largest stock exchange in Latin America (IMF (2016)) and was the sixth largest derivative market worldwide in 2015 (Statista). The most traded derivatives on BMF are US Dollar futures denominated in Brazilian Real (BRL). Due to the heavy regulation on the spot exchange market, Ventura and Garcia (2012) find that US Dollar futures are responsible for most of the price discovery in Brazilian Real, the most traded South American currency (BIS (2016)). CME is the world’s largest marketplace for derivatives and offers a Brazilian Real futures contract with almost identical conditions to the contract traded in BMF. The BRL futures in CME are denominated in US-Dollar and consequently, a long position in CME’s BRL futures is equivalent to a short position in BMF’s USD futures. Table 1 provides a short overview of the contracts. For both contracts, the last business day of each month is the last trading day for the contract expiring in the consecutive month. The contract expiring in the next month is generally the most traded contract in both markets and hence the contract I am focusing on at each point in time.

The trading volume of this contract in BMF is several times higher than in CME. Figure 9 shows the daily volume of traded contracts in BMF (blue line) and CME (black

<table>
<thead>
<tr>
<th>Location</th>
<th>CME (USA)</th>
<th>BMF (Brazil)</th>
</tr>
</thead>
<tbody>
<tr>
<td>End of trading</td>
<td>last business day of month</td>
<td>last business day of month</td>
</tr>
<tr>
<td>Contract size</td>
<td>100,000 BRL</td>
<td>50,000 USD</td>
</tr>
<tr>
<td>Quotation</td>
<td>USD per BRL</td>
<td>BRL per USD1,000.00</td>
</tr>
<tr>
<td>Tick size</td>
<td>USD 0.05 per BRL1,000.00</td>
<td>BRL 0.5 per USD1,000.00</td>
</tr>
<tr>
<td>Transaction cost</td>
<td>in USD</td>
<td>in USD</td>
</tr>
<tr>
<td>Trading hours</td>
<td>5pm-4pm CT</td>
<td>9am-6pm BRT</td>
</tr>
</tbody>
</table>
line) from November 2011 to February 2014. The dataset used in this paper does not provide quote data for the months between October 2012 and October 2013 and hence these months are excluded in the subsequent analysis. This is the area between the two dashed vertical lines. I am including this period in this graph for comparison purposes. The numbers are given in 100,000 contracts. Only days with trades in both markets are included and the last active day of each month is excluded.\textsuperscript{15} I find that the volume in CME in terms of contracts is on average 7\% of the BMF volume. There are 34 days where this number is at least 20\% and all of these days are between 02.03.2012 and 25.05.2012. This period is discussed further in Section 7. As shown in Figure 9, this increase is mostly due to spikes in CME activity. The red line provides the nominal exchange rate. Over these 28 months, the exchange rate was on average USD/BRL 2.06 and saw its minimum and maximum at 1.70 and 2.45, respectively. This implies that the contract size in BMF (USD 50,000) over this period was on average roughly equal to the CME contract in USD terms (BRL 100,000). The depreciation of the Brazilian Real over this period implies, however, that the size of the CME contract in USD terms decreases over time. At the end of the period, a single CME contract covers an amount equivalent to only USD 42,808.22.

Figure 9: Daily traded volume in BMF and CME

\textsuperscript{15} Trades outside BMF trading hours are also included, however, their total number is negligible.
trade in CME outside the BMF trading hours. CME is most active upon the opening of BMF, while BMF sees its peak in traded contracts at 11am. Both markets’ activity decreases over the course of the day with a local peak at 3pm. The last trading hour in BMF appear to see a stronger decline in BMF than CME.

Figure 10: Hourly traded volume in BMF and CME in percent

![Chart showing hourly traded volume in BMF and CME in percent.]

The black (blue) line illustrates the average number of traded contracts in 100,000 per hour in BMF (CME). Only days with trades in both markets are included. The last active day of each month is excluded.

An important property of the market for USD/BRA futures is that there are high costs for foreign firms to gain access to BMF. During the time considered in this paper, Brazilian law required non-resident investors to have a legal and fiscal representative present in Brazil. According to market participants in Brazil, CME’s market entry costs make it unattractive for most Brazilian investors to trade in CME directly. Additionally, over the horizon of our sample, the Brazilian government levies a transaction tax on international transactions and financial products. This Imposto sobre operações financeiras can serve as an additional deterrent for investors to be active in both markets. Together, this implies that the markets, while trading close to equivalent products, serve different sets of traders.

Again, based on the information from market participants, some traders serve as arbitrageurs connecting both markets while market makers, especially in CME, closely track movements in BMF in order to limit their exposure to arbitrageurs. This separation of the two markets is reflected in the theoretical model.

BMF data is directly provided by the exchange with millisecond precision. The dataset includes all trades and quotes. CME data is provided by Thomson Reuters and includes all trades as well as changes in the best bid and ask prices with millisecond


\[17\] It may also explain, why CME is able to provide a market in this product despite the far lower trading volume.
time stamps. The analysis focuses on the dynamics of the best bid and ask prices in both markets. Due to the inverse denomination of the contracts, a purchase of a CME contract is equivalent to the selling of a BMF contract. For the remainder of this paper, I invert the prices in CME in order to create time series with the same price denomination, meaning

\[ b_C = \frac{1}{\hat{a}_C} \quad \text{and} \quad a_C = \frac{1}{\hat{b}_C}, \tag{6.1} \]

where \( \hat{b}_C \) and \( \hat{a}_C \) are the best bid and ask prices denominated in USD for one BRL as reported by CME. The prices I use from here onwards are \( b_C \) and \( a_C \), the implied best bid and ask prices in BRL for one USD. I also adjusted the BMF prices in order to reflect the price for one USD instead of the contract quotation for USD 1,000.

The data horizon in this paper is from November 2011 to September 2012 and from November 2013 to February 2014 resulting in 248 trading days where both markets are open. The gap in the data horizon is due to data availability. The data contains observations for 9,817,200 seconds. 8,662,015.50 (88.2%) of these show positive spreads in both markets. \(^{18}99.82\% \) of the observations with positive spreads have a higher spread in CME than in BMF. The higher spread in CME indicates a lower liquidity in this venue most of the time. However, this does not necessarily imply that it is not sensible for liquidity or informed traders to open arbitrage opportunities in BMF as Section 7 shows.

Table 2: Spreads

<table>
<thead>
<tr>
<th></th>
<th>Seconds</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>9,817,200.00</td>
<td></td>
</tr>
<tr>
<td>Positive spread</td>
<td>8,662,015.50</td>
<td>88.23%</td>
</tr>
<tr>
<td>( S_C &gt; S_B )</td>
<td>8,646,173.87</td>
<td>99.82%</td>
</tr>
<tr>
<td>( S_C &lt; S_B )</td>
<td>15,841.63</td>
<td>0.18%</td>
</tr>
</tbody>
</table>

7 Empirics

The empirical Section of this paper consists of two parts. Firstly, I use the model in Section 3 in order to determine the effective transaction costs in the markets. I am using these transaction costs to locate the arbitrage opportunities and classify them into toxic and non-toxic arbitrage opportunities. In a second step, I calculate the toxic arbitrage information share and compare it to the Hasbrouck information share as well as the Gonzalo Granger permanent component share.

\(^{18}\)Negative spreads are a common problem when using high frequency data and observations with negative spreads are generally excluded from the analysis.
7.1 Transaction Costs

In order to identify the set of arbitrage opportunities, it is crucial to have a realistic estimate of the effective costs which an arbitrageur faces. While the trading venues CME and BMF provide transaction costs, these costs greatly vary depending on the type of investor. Further, it is necessary to allow for the possibility that traders will not engage in arbitrage without a minimum profit or without covering additional costs. In what follows, I am using an implication of the model in order to obtain an estimate of the effective transaction costs. According to the model, there is no reason for a size difference across types of arbitrage. The reason for this is that the size of all arbitrage opportunities is determined by $\gamma$, the difference in the mid-prices in both markets which is necessary for an arbitrage opportunity to be profitable. The derivation of this, can be found in the Appendix. Toxic arbitrage opportunities and non-toxic arbitrage opportunities are hence expected to have the same size. By size, I denote the difference between best bid and ask price across markets enabling arbitrage.

For an arbitrageur to break even, he has to cover the transaction costs in both markets.\(^{19}\) Further, the profit also needs to cover the transaction costs when liquidating his position in both markets. Hence, his total costs involve four individual transactions.\(^{20}\) The size of the arbitrage opportunity must exceed these costs

$$b_k - a_j > 2c_k + 2c_j \text{ where } k, j \in \{C, B\}, k \neq j,$$

where $c_k$ and $c_j$ are the transaction costs per Brazilian Real traded in market $k$ and $j$, respectively. I now take the implication from the model as a condition which has to hold given the correct transaction costs: Toxic and non-toxic arbitrage opportunities must on average be equal in size.\(^{21}\) I am choosing the lowest transaction cost for which this holds true. The estimation is under the assumption that the relevant costs for the arbitrageur have not changed over the time horizon. According to conversations with CME and BMF, this is true at least for the nominal transaction costs in both markets.

While the transaction costs do not change over time, the relative size of the contracts in both markets changes with the exchange rate as they are denominated in BRL and USD, respectively. I am correcting for this change in contract size as it also affects the effective transaction cost per USD in each market as nominal transaction costs are given per contract. A depreciation in the exchange rate implies that the contract in CME which covers BRL 100,000 is worth less in USD terms and therefore covers a different value compared to BMF which has its contract size fixed in USD. The condition for an

\(^{19}\)In the model, this is given by $\gamma = \frac{SB + SC}{2} + 2c$ as I am not assuming other costs than the transaction costs for the initial transactions. Once other costs arise e.g. for unravelling the positions, these costs will also be included in $\gamma$.

\(^{20}\)Say an arbitrageur buys in both markets at time $t = 1$. For these two trades, he needs to pay transaction costs. However, these traders also lead to a long position in each market. In order to close these positions when the price has returned to equilibrium at $t = 2$, he sells in both markets which again leads to transaction costs. There are four transactions in total, each leading to transaction cost.

\(^{21}\)This is a relaxed version of the implication in the model.
arbitrage opportunity hence becomes

\[ b_i - a_j > 2c_B + 2c_C \left( \frac{E_t}{2} \right) \]

where \( i,j \in \{C,B\}, i \neq j \),

\[(7.2)\]

where \( E_t \) is the exchange rate USD/BRL and \( \left( \frac{E_t}{2} \right) \) balances out the change in relative size of the contracts.

I find, that given transaction costs of on average USD 0.275 per contract in each market, there is no significant difference in size between toxic and non-toxic arbitrage opportunities. This number is in the range of transaction costs, which are given by the trading venues. In case of BMF this implies effective transaction costs of USD 5.50 per USD 1 million traded. I use this estimate of the transaction cost to identify the set of arbitrage opportunities. Table 3 shows the number of arbitrage opportunities and their total length. Using the transaction costs above, I find a total of 14,268 arbitrage opportunities with a total duration of 105,495.65 seconds or 29 hours and 18 minutes. This is equivalent to 1.22% of the sample. Of these 14,268 arbitrage opportunities, 98 are not used for the subsequent analysis as they either start or end with a negative spread or end with the end of the trading day. This leaves arbitrage opportunities of just over 29 hours. Two thirds of these arbitrage opportunities would be taken advantage of by executing sell orders in each market. I denote these opportunities as Type 1. The rest can be taken advantage of by executing buy orders in both markets, which I call Type 2 arbitrage. It is interesting to note that Type 2 arbitrage opportunities are not only more common but also appear to have a longer average duration as they make for 71% of the total duration of identifiable arbitrage opportunities.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Identifiable</th>
<th>Type 1 (Sell)</th>
<th>Type 2 (Buy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>14,268</td>
<td>14,170</td>
<td>9,250</td>
<td>4,920</td>
</tr>
<tr>
<td>Total time (sec)</td>
<td>105,495.65</td>
<td>104,690.27</td>
<td>74,188.97</td>
<td>30,501.30</td>
</tr>
<tr>
<td>% of pos. spread</td>
<td>1.22%</td>
<td>1.21%</td>
<td>0.86%</td>
<td>0.35%</td>
</tr>
</tbody>
</table>

Figure 11 shows the total number of arbitrage opportunities per week. 44% of these are found in March 2012. Five days (13th, 14th, 15th, 19th and 21st of March) have over 500 arbitrage opportunities each. These five days account for 38% of all arbitrage opportunities in our sample. It is worth noting that this is the same time where Figure 9 shows spikes in CME trading activity. This period is most likely to be explained by a change in the Imposto sobre operações financeiras (IOF) on March 15 2012. The IOF is a federal transaction tax in Brazil levied on credit, foreign exchange, insurance and securities transactions.

Table 4 shows the classification of the identified arbitrage opportunities using the chronology of price changes. CME/CME and BMF/BMF denote non-toxic arbitrage opportunities initiated in CME and BMF, respectively. Such arbitrage opportunities
are expected to be caused by liquidity shocks. CME/BMF and BMF/CME are toxic arbitrage opportunities caused by price relevant information in CME and BMF, respectively. As shown in Table 4, 1,011 (7.1%) of the arbitrage opportunities are toxic and initiated in CME. 5,976 (42.2%) are toxic arbitrage opportunities initiated in BMF. This implies that half of all arbitrage opportunities are toxic and most are caused by information arriving in BMF. The non-toxic arbitrage opportunities make 23.3% and 27.4%, respectively. There are 247 days in my sample which have at least one arbitrage opportunity. Table 5 provides summary statistics for these days.

Table 4: Arbitrage classification

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CME/CME</td>
<td>3,304</td>
<td>23.3%</td>
</tr>
<tr>
<td>CME/BMF</td>
<td>1,011</td>
<td>7.1%</td>
</tr>
<tr>
<td>BMF/CME</td>
<td>5,976</td>
<td>42.2%</td>
</tr>
<tr>
<td>BMF/BMF</td>
<td>3,879</td>
<td>27.4%</td>
</tr>
</tbody>
</table>

Figure 11: Weekly number of arbitrage opportunities
Table 5: Daily summary statistics

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full LOB CME</td>
<td>0</td>
<td>31,640</td>
<td>32,190</td>
<td>30,880</td>
<td>32,380</td>
<td>32,400</td>
</tr>
<tr>
<td>Full LOB BMF</td>
<td>7.66</td>
<td>32,350</td>
<td>32,350</td>
<td>31,340</td>
<td>32,360</td>
<td>32,400</td>
</tr>
<tr>
<td>Mean Spread CME</td>
<td>0.0012</td>
<td>0.0019</td>
<td>0.0027</td>
<td>0.0055</td>
<td>0.0040</td>
<td>0.2002</td>
</tr>
<tr>
<td>Mean Spread BMF</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0021</td>
</tr>
<tr>
<td>Contr. Traded CME</td>
<td>0</td>
<td>4,966</td>
<td>9,876</td>
<td>26,110</td>
<td>23,690</td>
<td>372,200</td>
</tr>
<tr>
<td>Contr. Traded BMF</td>
<td>13,520</td>
<td>195,200</td>
<td>254,900</td>
<td>250,200</td>
<td>301,800</td>
<td>496,000</td>
</tr>
<tr>
<td>Arb. Duration</td>
<td>0.01</td>
<td>11.32</td>
<td>112.60</td>
<td>423.80</td>
<td>303.50</td>
<td>20,010.00</td>
</tr>
<tr>
<td>Number of Arbs.</td>
<td>1</td>
<td>5</td>
<td>20</td>
<td>57.37</td>
<td>44.50</td>
<td>1,457.00</td>
</tr>
<tr>
<td>CME/BMF</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>4.09</td>
<td>1.00</td>
<td>297.00</td>
</tr>
<tr>
<td>BMF/CME</td>
<td>0.00</td>
<td>2.00</td>
<td>5.00</td>
<td>24.19</td>
<td>15.00</td>
<td>728.00</td>
</tr>
<tr>
<td>CME/CME</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>13.38</td>
<td>4.50</td>
<td>644.00</td>
</tr>
<tr>
<td>BMF/BMF</td>
<td>0.00</td>
<td>2.00</td>
<td>8.00</td>
<td>15.70</td>
<td>19.00</td>
<td>206.00</td>
</tr>
</tbody>
</table>

The table displays summary statistics for days with at least one arbitrage opportunity. Full LOB CME and Full LOB BMF is the number of seconds in a day with a full limit order book in CME and BMF, respectively. A full limit order book means, that there are both best bid and ask prices available. Mean Spread is the mean value of the spread in each market. Contr. Traded is the number of traded contracts in each market. Arb. Duration is the total duration of arbitrage opportunities in a day in seconds.

There are five days which see a full limit order book for less than 8 hours (28,800 seconds), two of these are due to CME and three are due to BMF. The median time without a full limit order book in CME is 3.5 minutes per day. In BMF the median time is less than a minute. As indicated before, the mean spread in CME is consistently higher than in BMF. The minimum number of trades in CME is none compared to 13,520 traded contracts in BMF. The total duration of arbitrage opportunities in one day varies between 0.01 seconds and 20,010 seconds (5 hours and 33 minutes). The total number of arbitrage opportunities for days with at least one arbitrage opportunity varies between 1 and 1,457. 229 out of 247 days see at least one toxic arbitrage opportunity. On 228 out of 247 days, there is at least one non-toxic arbitrage opportunity.

7.2 Toxic Arbitrage Information Share

The summary statistics of the arbitrage opportunities in Table 4 show that most of the toxic arbitrage opportunities (5,976 compared to 1,011) are caused by information arriving in BMF rather than in CME. However, as highlighted in Section 3, the frequency of toxic arbitrage opportunities alone does not provide an unbiased measure of information arrival. In order to calculate the unbiased information share, I require measures for the right-hand-side parameters of Equation (4.20): $IS^B_{	ext{log}_e}$, $\Psi$ and $\Gamma$.

When using the distribution of price impacts for both $B/C$ and $C/B$ arbitrage opportunities, it is important to use the impact of the two information events on the same market, i.e. either on the price series of BMF or CME. I use the impact on BMF mid-prices here, due to BMF’s higher liquidity. For the transaction cost, I use the same estimate as in the determination of the arbitrage opportunities, i.e. USD 5.50 per USD.
1 million. It is worth noting that in this application the effect of the transaction costs on \( T\text{AIS} \) appears to be marginal.

In what follows, I am focusing on days with at least ten toxic arbitrage opportunities in order to keep the error around the toxic arbitrage information share sufficiently small. This leads to 106 days or over 42\% of the total number of days with at least one arbitrage opportunity.\(^{22}\) The results do not change when including all observations. The spreads in both markets, \( S_C \) and \( S_B \), are observable as long as open bid and open ask orders exist. As before, I follow the procedure as introduced in Section 4.

The summary statistics for the mean spreads are given in the first two rows of Table 6. \( IS_{\text{tox}}^B \) in row three provides the summary statistics for the ratio of toxic arbitrage opportunities. The mean value of close to 90\% already suggests a much higher share of information arriving in BMF. However, as highlighted before, this measure is biased due to the different price impact of information in both markets and the different spreads.

Following the numerical approximation in Equation (4.9) for both BMF and CME and solving for the standard deviations of the price impacts of arriving information leads to \( \sigma_B = 0.00031 \) and \( \sigma_C = 0.00024 \), respectively. Hence, the information arriving in BMF has a larger price impact than information arriving in CME. Together with the lower spreads in BMF, this would suggest that the percentage of BMF initiated toxic arbitrage opportunities \( IS_{\text{tox}}^B \) is underestimating the informational share of BMF. Using the approximated \( \sigma_B \) and \( \sigma_C \), I am now able to estimate the daily \( \Psi \).

Using \( IS_{\text{tox}}^B \), daily \( \Psi \) as well as \( \sigma_B \) and \( \sigma_C \), I am able to determine the bias given by the difference between the toxic arbitrage ratio and the unbiased information share. The bias is presented in the fourth row. As suspected, the toxic arbitrage information share is consistently higher than \( IS_{\text{tox}}^B \). The mean bias of -0.09 implies that the ratio of toxic arbitrage opportunities underestimates the percentage of information entering BMF by 9 percentage points on average. The bias ranges from an underestimation of 72.3 percentage points to an unbiased estimation. Row five provides the resulting unbiased information share \( TAIS \) as given by Equation (4.20). I find that the unbiased information share for BMF is between 64.7\% and 100\%. This means, that between 64.7\% and 100\% of price relevant information arrives first in BMF. The median toxic arbitrage information share is 99.8\% while the mean is 99\%. This is very much in line with the information I received from conversations with market participants in Brazil who expect no gain in information from observing the prices in CME.

The unbiased information share based on toxic arbitrage opportunities provides an unbiased, model based measure of information arriving in each market. It is reasonable to expect this to be similar in size as other measures of price discovery. The standard procedures for price discovery are given by the information share (\( IS \)) introduced by Hasbrouck (1995) and the permanent component share (\( CS \)) by Gonzalo and Granger (1995). For a detailed description and implementation of these procedures see e.g. Tse et al. (2006), Chen and Gau (2010) and Yan and Zivot (2010).

In the remainder of this section, I first look at the summary statistics of the information shares in order to compare their overall properties. In a second step, I compare

\[^{22}\text{This is equivalent to 35\% of the total sample.}\]
Table 6: Unbiased information share based on toxic arbitrage (#tox ≥ 10)

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spread CME</td>
<td>0.00015</td>
<td>0.00024</td>
<td>0.00030</td>
<td>0.00040</td>
<td>0.00045</td>
<td>0.00270</td>
</tr>
<tr>
<td>Spread BMF</td>
<td>0.00050</td>
<td>0.00050</td>
<td>0.00050</td>
<td>0.00056</td>
<td>0.00050</td>
<td>0.00250</td>
</tr>
<tr>
<td>$IS_{B_{tox}}$</td>
<td>0.1600</td>
<td>0.8648</td>
<td>0.9444</td>
<td>0.8961</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.7233</td>
<td>-0.1293</td>
<td>-0.0934</td>
<td>-0.0545</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>TAIS</td>
<td>0.6470</td>
<td>0.9934</td>
<td>0.9981</td>
<td>0.9895</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

the behaviour of these measures over time in order to get a better idea of their stability and persistency.

The first row of Table 7 again provides the summary statistics for the toxic arbitrage information share (TAIS). Rows two and three provide the summary of the permanent component share (CS) and the Hasbrouck information share (IS), respectively. Following Hasbrouck (1995, 2003), I use secondly frequency of best bid and offer prices in order to estimate the daily component and information shares. The toxic arbitrage information share is close to the Gonzalo Granger component share. The difference between the two measures in median and mean is 2.8 and 7.5 percentage points, respectively. When using the procedure introduced by Hasbrouck (1995), the information share depends on the ordering of the variables and can differ substantially. In the two variable case, the two possible orderings provide an upper and a lower bound. Hasbrouck (2002) shows that all permutations need to be taken into account. Row three provides the midpoint of the two information shares resulting from the Hasbrouck procedure. The Hasbrouck information share is lower than the permanent component share of Gonzalo Granger.

Table 7: Comparison of Information shares (#tox ≥ 10)

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAIS</td>
<td>0.6470</td>
<td>0.9934</td>
<td>0.9981</td>
<td>0.9895</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>CS</td>
<td>0.2736</td>
<td>0.8987</td>
<td>0.9703</td>
<td>0.9146</td>
<td>0.9939</td>
<td>1.0000</td>
</tr>
<tr>
<td>IS</td>
<td>0.0192</td>
<td>0.7858</td>
<td>0.9607</td>
<td>0.8612</td>
<td>0.9876</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The table shows the summary statistics of the toxic arbitrage information share (TAIS), permanent component share (CS) and Hasbrouck information share (IS) for days with at least 10 toxic arbitrage opportunities. TAIS are based on centisecondly frequency. CS and IS are based on secondly frequency.

23While Hasbrouck (1995) uses best quotes as in the application here, some studies use average transaction price in given intervals, generally between 1 second and 5 minutes. Even when using 5 minute intervals, there will not be a single trade in around 50% of the observations for CME. Following this procedure leads to $IS$ and $CS$ price discovery shares with unrealistically high information shares for CME.
The summary statistics give an impression of how the price discovery measures relate to each other. Figures 12a and 12b illustrate the development of these shares over time. As in Table 6, I am only looking at days with at least ten toxic arbitrage opportunities. Figure 12a illustrates the evolution of $TAIS$ (red line) and the permanent component share (blue line). The black vertical line marks the data gap in my sample between October 2012 and October 2013. I find that both measures are close to 100% most of the time. The component share is more volatile and shows several instances where CME has a larger price discovery share than BMF. There is only one instance where the toxic arbitrage information share is below 80%. At this time, the permanent component share also shows a strong drop, though significantly more extreme.

Figure 12: Price discovery over time

![Graph showing price discovery over time](image)

- a) Permanent component share
- b) Hasbrouck information share

The red line in Figure 12b again illustrates the toxic arbitrage based information share for days with at least ten toxic arbitrage opportunities. The green line shows the midpoint of the Hasbrouck information share ($IS$). The green area illustrates the range between the upper and the lower estimate from the Hasbrouck procedure. While the upper bound remains close to 100% most of the time, I still find that there are several instances where both bounds drop far below the level of the unbiased information share. While Hasbrouck always shows a drop when Gonzalo Granger does, there are several more outliers for Hasbrouck.

In summary, all three measures for price discovery clearly show a higher share for BMF. However, the toxic arbitrage information share is more persistent, less volatile and overall more in line with market participants expectations. When comparing these results with the simulations in Section 5, we find the following. Whenever, simulating data such that the more liquid market has an information share above 80%, the toxic arbitrage information share has rather underestimated the actual information share than overestimated. Furthermore, simulation III has shown that missing observations in the bid and ask let to a much better performance of the $TAIS$ compared to the alternatives.
for information share above 80%. These results indicate that the TAIS provides a more realistic estimate in this setting.

Overall, I find evidence that the toxic arbitrage information share can be a valuable alternative as it performs better in less liquid markets.

8 Conclusion

This paper introduces a model to analyse the relationship of information arrival and toxic arbitrage opportunities in a two market setting. Using this model, I derive a measure of information share which is based on the occurrence of toxic arbitrage opportunities and corrects for distortions due to differences in the informational structure, the spreads and the transaction costs in both markets. In order to test the properties of the toxic arbitrage based information share, I run a set of simulations as well as using a unique dataset of US Dollar/Brazilian Real futures traded in Chicago and São Paulo. Both the simulations as well as the dataset allow me to evaluate the performance of the toxic arbitrage information share in a market setting with low liquidity in one of the markets, namely CME.

The simulations as well as the application show that the estimates based on the toxic arbitrage information share are more persistent and less volatile than the standard procedures by Hasbrouck (1995) and Gonzalo and Granger (1995). This is due to the higher robustness of the toxic arbitrage information share to low liquidity and partially empty limit order books. It is therefore a crucial alternative given the growing fragmentation of financial markets. The median estimate for the toxic arbitrage information share is close to the standard price discovery share. The difference is mostly driven by downward spikes in the latter. Generally, the results for the toxic arbitrage information share are more in line with the opinions of market participants in BMF and CME than the standard procedures.

In summary, my results suggest that the toxic arbitrage information share is a good addition to the analysis of information shares and price discovery especially when involving low liquidity markets. Being derived in complete separation of standard approaches, the toxic arbitrage information share avoids most of the problems of using VAR based information shares. Due to the binary nature of toxic arbitrage opportunities, the model suggests that it is also possible to create error bands for the toxic arbitrage information share. The exploration of this will be part of future work.
References


38


A Appendix

A.1 Solution of Model

The following provides the complete solution of the model presented in this paper. There are two markets $B$ and $C$ trading the same asset, three dates ($t \in \{0, 1, 2\}$), two market makers, one in each market, an arbitrageur between markets and both a liquidity and an informed trader in each market. In total, there are seven agents. The central point of this setup is that due to market entry costs the arbitrageur is the only market participant able and willing to trade in both markets and her orders move faster than information on the prices in each market. Market $C$ can be thought of as CME and market $B$ as BMF. Figure 13 provides the time line of the model. The description below is in reverse order.

Figure 13: Time Line

<table>
<thead>
<tr>
<th>$t = 0$</th>
<th>$t = 1$</th>
<th>$t = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrival of informed or liquidity trader</td>
<td>Posting of Prices and Trades</td>
<td>Realization of the Final Value $\theta$</td>
</tr>
</tbody>
</table>

At $t = 2$ the final value $\theta = \mu + \epsilon$ of the asset is realised. The final value of the asset is either high or low

$$\theta = \begin{cases} 
\mu + \sigma_n & \text{with probability } 0.5 \\
\mu - \sigma_n & \text{with probability } 0.5 
\end{cases}$$

where $\sigma_n$ is the variation in the fundamental value of the asset. The expected value of the asset is hence $E(\theta) = \mu + E(\epsilon) = \mu$. Market maker $j \in \{B, C\}$ is specialised in market $j$ and is only active in this market. Due to the physical distance between the market places, the information is not instantaneously available in both markets possibly leading to short lived arbitrage opportunities.

At $t = 1$, market makers simultaneously post ask price $a_j$ and bid price $b_j$ for $j \in \{B, C\}$ such that:

$$a_j = v_j + \frac{S_j}{2}, \quad \text{and} \quad b_j = v_j - \frac{S_j}{2}$$

where $v_j$ is market maker $j$’s valuation for the asset and $S_j$ is his bid-ask spread. Quotes are for a fixed quantity, normalised to 1. Market maker $j$’s valuation $v_j$ is determined at time $t = 0$. This valuation differs for the following four events:

- Liquidity trade in market $B$
- Liquidity trade in market $C$

\[40\]
• News arrival in market B
• News arrival in market C

Figure 14 provides an illustration of the probabilities of each event. In what follows, I develop each of these cases separately.

![Figure 14: Cases and Probabilities](image)

The image illustrates the probability of each case. *Liq-trade* stands for liquidity trade in either market B or C.

The liquidity trader derives a utility of $\mu + \epsilon_L$ from the asset at time $t = 2$, where $\epsilon_L = \sigma_L$ or $\epsilon_L = -\sigma_L$ with equal probability. In the first case, the liquidity trader has a positive liquidity need, i.e. he wants to buy the asset. Here, $\sigma_L$ can be thought of as the added willingness to pay due to the liquidity need. It is assumed to be large enough to allow for a profitable trade in the first place. If this is not given, nothing happens. In the case of $\epsilon_L = -\sigma_L$, the liquidity trader wants to sell the asset. The probability that a liquidity trader arrives at one of the markets is $(1 - \alpha)$.

The market makers B and C derive the expected utility $\theta$ from the asset owned at time $t = 2$. Their valuation is hence $v_B = v_C = E(\theta) = \mu$. Given that a liquidity trader arrives, the probability that this trader arrives in market B is given by $\rho_b$. This implies that $\rho_b$ is the fraction of liquidity traders arriving in market B and $(1 - \rho_b)$ is the fraction of liquidity traders in market C.

Next, consider the case where a liquidity trader arrives in market B. The probability for this happening is given by $(1 - \alpha)\rho_b$. Say, the liquidity trader has a positive demand for the asset, i.e. $\epsilon_L = \sigma_L$. She has two options. The liquidity trader can execute a market order (case $\kappa_B = M$) against market maker B’s best ask or post a limit order ($\kappa_B = L$). If she uses a market order, her and the market maker’s profit are given by

$$
\text{Profit}_{\text{Liq}} = \mu + \sigma_L - a_B - c = \mu + \sigma_L - \left(\mu + \frac{S_B}{2}\right) - c = \sigma_L - \frac{S_B}{2} - c, \quad (A.3)
$$

$$
\text{Profit}_{\text{MM}} = a_B - \mu = \mu + \frac{S_B}{2} - \mu = \frac{S_B}{2}, \quad (A.4)
$$

\[24\]The probability that the liquidity trader arrives and is willing to trade will depend on the spread as well as the transaction costs. This is simplified here as the focus of the paper is on toxic arbitrage and the main results will not be affected by this.
where $c$ is the transaction cost in each market.\footnote{The transaction costs apply to market orders and executed limit orders by liquidity and informed traders. Market makers do not face transaction costs. It follows that if no market maker is involved in a trade, both sides will have to pay the transaction costs. These costs cover the infrastructure of the exchange and do not go to the market maker.} In the second scenario, $\kappa_B = L$, the liquidity trader will post a limit order of $b_B^T = \mu - \frac{S_B}{2} + \gamma$, where $\gamma > 0$ is endogenously chosen by the liquidity trader. It is the price improvement the liquidity trader needs to offer compared to the market maker’s best price in order to incentivise the arbitrageur to trade. The size of $\gamma$ is derived below. Given that $b_B^T = b_B + \gamma > b_B$, it follows that $b_B^T$ is now the best bid in market $B$. The idea is to post a limit order which will be picked up by the arbitrageur. The arbitrageur will buy at this price if he can make a profit by simultaneously selling the asset in market $C$. Due to the trades in both markets, the arbitrageur incurs a total cost of $2c$. The profit of the arbitrageur is then given by

$$Profit_{Arb} = b_B^T - a_C - 2c$$ \hspace{1cm} (A.5)

$$= \left( \mu - \frac{S_B}{2} + \gamma \right) - \left( \mu + \frac{S_C}{2} \right) - 2c$$ \hspace{1cm} (A.6)

$$= \gamma - \frac{S_C + S_B}{2} - 2c.$$ \hspace{1cm} (A.7)

The liquidity trader will choose the smallest $\gamma$ which still leads the arbitrageur to engage in arbitrage, i.e.

$$\gamma = \frac{S_C + S_B}{2} + 2c,$$ \hspace{1cm} (A.8)

leading to $Profit_{Arb} = 0$. In this case, the liquidity trader uses the arbitrageur as a channel to trade against market maker $C$. Hence, $\gamma$ can be thought of as the price of the arbitrageurs’ service of providing a channel for indirectly trading against the market maker in the other market. Market maker $C$ receives a profit of

$$Profit_{MMB} = a_C - \mu = \mu + \frac{S_C}{2} - \mu = \frac{S_C}{2}.$$ \hspace{1cm} (A.9)

In this case, the liquidity traders profit is given by

$$Profit_{LiqL} = \mu + \sigma_L - b_B^T - c = \mu + \sigma_L - \left( \mu - \frac{S_B}{2} + \gamma \right) - c$$

$$= \sigma_L + \frac{S_B}{2} - \left( \frac{S_C + S_B}{2} + 2c \right) - c = \sigma_L - \frac{S_C}{2} - 3c.$$ \hspace{1cm} (A.10)

His profit includes the transaction costs of his own and indirectly the arbitrageur’s trades. Comparing the two options expressed in Equations (A.3) and (A.10), I find that the liquidity trader will post a market order if

$$\sigma_L - \frac{S_B}{2} - c \geq \sigma_L - \frac{S_C}{2} - 3c$$

$$\Rightarrow \frac{S_C}{2} - \frac{S_B}{2} \geq -2c.$$ \hspace{1cm} (A.11)
This implies that despite the fact that each liquidity trader has access to only one of the two markets, he still has the option of trading indirectly against the other market maker by incentivising the arbitrageur. Rather than being the driving force of arbitrage as in Foucault et al. (2016), the arbitrageur becomes a tool used by the liquidity trader.

With probability \((1 - \alpha)(1 - \rho_b)\), a liquidity trader arrives at market \(C\). As before, she needs to choose between two cases, \(\kappa_C = M\) and \(\kappa_C = L\). With the same reasoning as above, the liquidity trader will post a market order in market \(C\), if

\[
\sigma_L - \frac{S_C}{2} - c \geq \sigma_L + \frac{S_C}{2} - \gamma - c \tag{A.12}
\]

\[\Rightarrow 2c \geq \frac{S_C - S_B}{2}, \tag{A.13}\]

leading to a profit of \(\frac{S_C}{2}\) for market maker \(C\). If \(2c < \frac{S_C - S_B}{2}\), the liquidity trader will post a limit order in market \(C\). When trying to sell the asset, the liquidity trader will post an order with a bid price of \(b^T_C = \mu - \frac{S_C}{2} + \gamma\). Given that \(b^T_C > b_C\), this is now the best bid in market \(C\). The trader will choose the smallest \(\gamma\) which still opens a profitable arbitrage opportunity i.e. \(b^T_C \geq a_B + 2c\). It follows that \(\gamma\) is again given by Equation (A.8). The arbitrageur then buys in market \(C\) and sells in market \(B\), leading to a profit of \(\gamma - \frac{S_C + S_B}{2} - 2c = 0\). This time, the market maker in \(B\) receives a profit of \(\frac{S_B}{2}\). Cases where the liquidity trader posts a limit order lead to what Foucault et al. (2016) describe as non-toxic arbitrage opportunities. These arbitrage opportunities have a positive effect on welfare as they help to balance liquidity across markets.

With probability \(\alpha\), an informed trader arrives at one of the markets. The informed trader observes \(\epsilon\), the change in the fundamental value as shown in Equation (A.1) at time \(t = 1\). Hence, he uses this information to trade before the information becomes public knowledge at time \(t = 2\). This news arrival via the informed trader leads to what Foucault et al. (2016) describe as toxic arbitrage. Toxic arbitrage opportunities are not considered welfare improving. The arbitrageur uses his lower latency in order to trade against stale trades by the other market maker. The informed trader arrives at market \(B\) with probability \(\alpha \delta_b\). Say the informed trader receives negative information about the asset, i.e. the informed trader observes \(\epsilon = -\sigma_n\). Hence, the true value of the asset is given by \(\theta = \mu - \sigma_n\). He will use this information in the two ways which were also available for the liquidity trader. However, in contrast to the liquidity trader who only has a limited demand of one unit, the informed trader will trade the maximum amount possible to take advantage of his information. Rather than choosing between limit and market order, he uses both as long as it is profitable to do. The maximum possible traded volume in this model is given by two units; one in each market.

If the anticipated price impact \(\sigma_n\) is sufficiently high, the following holds. The informed trader executes a sell market order of one unit in market \(B\) in order to trade
against the market maker at price $b_B = \mu - \frac{S_B}{2}$, leading to profits of

\[
\text{Profit}_{Inf_M} = b_B - (\mu - \sigma_n) - c = \left( \mu - \frac{S_B}{2} \right) - (\mu - \sigma_n) - c \tag{A.14}
\]

\[
= \sigma_n - \frac{S_B}{2} - c, \tag{A.15}
\]

\[
\text{Profit}_{MM_B} = (\mu - \sigma_n) - b_B = \mu - \sigma_n - \left( \mu - \frac{S_B}{2} \right) \tag{A.16}
\]

\[
= - \left( \sigma_n - \frac{S_B}{2} \right). \tag{A.17}
\]

At the same time, he will post a sell limit order of one unit at price $a_T^B = \mu + \frac{S_B}{2} - \gamma$ in order to be picked up by the arbitrageur. Again, $\gamma$ will be chosen as $\gamma = \frac{S_C + S_B}{2} + 2c$ leading to zero profit for the arbitrageur. This results in additional profits of

\[
\text{Profit}_{Arb} = b_C - a_T^B - 2c = \left( \mu - \frac{S_C}{2} \right) - \left( \mu + \frac{S_B}{2} - \gamma \right) - 2c \tag{A.18}
\]

\[
= - \frac{S_C + S_B}{2} + \gamma - 2c = 0, \tag{A.19}
\]

\[
\text{Profit}_{Inf_L} = a_T^B - (\mu - \sigma_n) - c \tag{A.20}
\]

\[
= \left( \mu + \frac{S_B}{2} - \gamma \right) - (\mu - \sigma_n) - c \tag{A.21}
\]

\[
= \sigma_n - \frac{S_C}{2} - 3c, \tag{A.22}
\]

\[
\text{Profit}_{MM_C} = (\mu - \sigma_n) - b_C = (\mu - \sigma_n) - \left( \mu - \frac{S_C}{2} \right) \tag{A.23}
\]

\[
= - \left( \sigma_n - \frac{S_C}{2} \right). \tag{A.24}
\]

Each market maker $j$ consequently incurs a loss of $- \left( \sigma_n - \frac{S_j}{2} \right)$. The informed trader is able to trade a total amount of two units. He is trading directly against market maker $B$ and uses the arbitrageur to trade indirectly against market maker $C$. Combining the two sub-profits of the informed trader in Equations (A.15) and (A.22) leads to a total profit of

\[
\text{Profit}_{Inf} = \text{Profit}_{Inf_M} + \text{Profit}_{Inf_L}
\]

\[
= \left( \sigma_n - \frac{S_B}{2} - c \right) + \left( \sigma_n - \frac{S_C}{2} - 3c \right)
\]

\[
= 2\sigma_n - \frac{S_B + S_C}{2} - 4c \tag{A.25}
\]

The probability that the news first arrive in market $C$ is given by $\alpha(1 - \delta_b)$. Again, the informed trader will use market and limit orders to take advantage of the information,
leading to the same profits for each market maker, the arbitrageur and the informed trader as in the previous case.

Table 8 shows the market participants’ payoffs in detail under the assumption that \( \sigma_n \) is sufficiently large. The implication is that it is optimal for the informed trader to use both limit and market orders. For each of the market participants and each case, Table 8 shows the gain, loss and net payoff including and excluding transaction costs. Gain refers to the value of the bought asset or the price earned by selling the asset. Loss refers to the price payed when buying the asset or the value of a sold asset. Net Payoff is given by Gain minus Loss. The Profit is given by the Net Payoff minus transaction costs. For Market makers there are no transaction costs and hence the Net Payoff is equal to the Profit. The last line of Table 8 highlights which market initiates and closes the deviation in prices. Toxic arbitrage opportunities are initiated by a price movement in one market and closed by a price movement in the other market. Non-toxic arbitrage opportunities are closed in the same market as they are initiated once the liquidity demand is met. I denote by \( \kappa_C \in \{M, L\} \) if condition \( 2c \geq \frac{S_C - S_B}{2} \) as stated in Equation (A.13) is met or not. If \( \kappa_C = M \), the liquidity or informed trader will post market orders if he arrives in market \( C \). Similarly, \( \kappa_B \in \{M, L\} \) is \( M \) if condition \( \frac{S_C - S_B}{2} \leq -2c \) as given in Equation (A.11) is met, i.e. if it is optimal for the liquidity or informed trader arriving in market \( B \) to post a market order.
Table 8: Trader’s expected payoffs

<table>
<thead>
<tr>
<th></th>
<th>news in $B$</th>
<th>news in $C$</th>
<th>liq. trade $B$ ($\kappa_B = M$)</th>
<th>liq. trade $B$ ($\kappa_C = L$)</th>
<th>liq. trade $C$ ($\kappa_C = M$)</th>
<th>liq. trade $C$ ($\kappa_C = L$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit</td>
<td>$\alpha_0$</td>
<td>$\alpha(1 - \delta_B)$</td>
<td>$(1 - \alpha)\rho_B$</td>
<td>$(1 - \alpha)\rho_B$</td>
<td>$(1 - \alpha)(1 - \rho_B)$</td>
<td>$(1 - \alpha)(1 - \rho_B)$</td>
</tr>
<tr>
<td>Market maker $C$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain</td>
<td>$\mu - \sigma_n$</td>
<td>$\mu - \sigma_n$</td>
<td>$\mu + \sigma_C$</td>
<td>$\mu + \sigma_C$</td>
<td>$\mu + \sigma_C$</td>
<td>$\mu + \sigma_C$</td>
</tr>
<tr>
<td>Loss</td>
<td>$\mu - \sigma_C$</td>
<td>$\mu - \sigma_C$</td>
<td>$\mu$</td>
<td>$\mu$</td>
<td>$\mu$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>Profit</td>
<td>$\left(\sigma_n - \frac{\rho_C}{\delta_C}\right)$</td>
<td>$\left(\sigma_n - \frac{\rho_C}{\delta_C}\right)$</td>
<td>$\frac{\rho_C}{\delta_C}$</td>
<td>$\frac{\rho_C}{\delta_C}$</td>
<td>$\frac{\rho_C}{\delta_C}$</td>
<td>$\frac{\rho_C}{\delta_C}$</td>
</tr>
<tr>
<td>Market maker $B$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain</td>
<td>$\mu - \sigma_n$</td>
<td>$\mu - \sigma_n$</td>
<td>$\mu + \sigma_B$</td>
<td>$\mu + \sigma_B$</td>
<td>$\mu + \sigma_B$</td>
<td>$\mu + \sigma_B$</td>
</tr>
<tr>
<td>Loss</td>
<td>$\mu - \sigma_B$</td>
<td>$\mu - \sigma_B$</td>
<td>$\mu$</td>
<td>$\mu$</td>
<td>$\mu$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>Profit</td>
<td>$\left(\sigma_n - \frac{\rho_B}{\delta_B}\right)$</td>
<td>$\left(\sigma_n - \frac{\rho_B}{\delta_B}\right)$</td>
<td>$\frac{\rho_B}{\delta_B}$</td>
<td>$\frac{\rho_B}{\delta_B}$</td>
<td>$\frac{\rho_B}{\delta_B}$</td>
<td>$\frac{\rho_B}{\delta_B}$</td>
</tr>
<tr>
<td>Liquidity/informed trader</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain</td>
<td>$2\mu - \gamma$</td>
<td>$2\mu - \gamma$</td>
<td>$\mu + \sigma_L$</td>
<td>$\mu + \sigma_L$</td>
<td>$\mu + \sigma_L$</td>
<td>$\mu + \sigma_L$</td>
</tr>
<tr>
<td>Loss</td>
<td>$2\mu - 2\sigma_n$</td>
<td>$2\mu - 2\sigma_n$</td>
<td>$\mu + \frac{\sigma_B}{\gamma}$</td>
<td>$\mu + \frac{\sigma_B}{\gamma}$</td>
<td>$\mu + \frac{\sigma_B}{\gamma}$</td>
<td>$\mu + \frac{\sigma_B}{\gamma}$</td>
</tr>
<tr>
<td>Net Payoff</td>
<td>$2\sigma_n - \gamma$</td>
<td>$2\sigma_n - \gamma$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma}$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma}$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma}$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma}$</td>
</tr>
<tr>
<td>Profit</td>
<td>$2\sigma_n - \gamma - 2c$</td>
<td>$2\sigma_n - \gamma - 2c$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma} - c$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma} - c$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma} - c$</td>
<td>$\sigma_L - \frac{\sigma_B}{\gamma} - c$</td>
</tr>
<tr>
<td>Arbitrageur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain</td>
<td>$\mu - \frac{\sigma_C}{\gamma}$</td>
<td>$\mu - \frac{\sigma_B}{\gamma}$</td>
<td>$\mu - \frac{\sigma_B}{\gamma}$</td>
<td>$\mu - \frac{\sigma_B}{\gamma}$</td>
<td>$\mu - \frac{\sigma_B}{\gamma}$</td>
<td>$\mu - \frac{\sigma_B}{\gamma}$</td>
</tr>
<tr>
<td>Loss</td>
<td>$\mu + \frac{\sigma_B}{\gamma} + \gamma$</td>
<td>$\mu + \frac{\sigma_B}{\gamma} + \gamma$</td>
<td>$\mu + \frac{\sigma_B}{\gamma} + \gamma$</td>
<td>$\mu + \frac{\sigma_B}{\gamma} + \gamma$</td>
<td>$\mu + \frac{\sigma_B}{\gamma} + \gamma$</td>
<td>$\mu + \frac{\sigma_B}{\gamma} + \gamma$</td>
</tr>
<tr>
<td>Net Payoff</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c$</td>
</tr>
<tr>
<td>Profit</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c - 2c$</td>
<td>$\gamma - \frac{\sigma_B}{\gamma} - 2c - 2c$</td>
</tr>
</tbody>
</table>

The payoffs shown here are each for either positive liquidity shocks or negative information shocks. The results are symmetric, however. Gain refers to the value of the bought asset or the price earned by selling the asset. Loss refers to the price payed when buying the asset or the value of a sold asset. Net Payoff is given by Gain minus Loss. The Profit is given by the Net Payoff minus transaction costs. There are no transaction costs for market makers and hence the Net Payoff is equal to the Profit. $\gamma$ in equilibrium is given by $\gamma = \frac{\sigma_B}{\gamma} + 2c$ leading to zero profits for the arbitrageurs.
Table 9 illustrates the conditions for the combinations of \( \kappa_C \) and \( \kappa_B \). Given non-negative transaction costs, it becomes clear that the combination \( \kappa_C = L, \kappa_B = L \) is not feasible. It cannot be lucrative to post limit orders in both markets in order to get picked up by the arbitrageur. In at least one of the markets, a market order must be optimal. The reason for this is that each transaction incurs transaction costs. Limit orders followed by arbitrage trades in both markets involve more trades than a market order and hence incur more costs. A sufficiently large spread in one market compared to the other can still justify the increase in transaction costs and therefore a limit order. However, it is not feasible that limit orders are optimal in both directions. This reduces the number of possible equilibria to 3. The markets can only be in one of the three equilibria at a time and the spreads in both markets determine the equilibrium.

Table 9: Scenario Differentiation

<table>
<thead>
<tr>
<th>( \kappa_C )</th>
<th>( \kappa_B = M )</th>
<th>( \kappa_B = L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa_C = M )</td>
<td>( 2c \geq \frac{S_C - S_B}{2} \geq -2c )</td>
<td>( 2c &gt; \frac{S_C - S_B}{2} &gt; 2c )</td>
</tr>
<tr>
<td>( \kappa_C = L )</td>
<td>( \frac{S_C - S_B}{2} &gt; 2c )</td>
<td>( -2c &gt; \frac{S_C - S_B}{2} &gt; 2c )</td>
</tr>
</tbody>
</table>

### A.2 Size of Arbitrage Opportunities

Toxic arbitrage opportunities can occur in each of these equilibria. In contrast, non-toxic arbitrage opportunities can only occur if it is optimal for the liquidity trader in one of the markets to post a limit order. Independently of the type of equilibrium, the size of realised arbitrage opportunities is the same for both toxic and non-toxic arbitrage opportunities. I denote by the size of an arbitrage opportunity the difference between the best bid in one market and the best ask in the other. As highlighted in the model, arbitrage opportunities occur if a limit order is posted by either a liquidity or informed trader in order to incentivise the arbitrageur to trade. Such an ask limit order needs to be at price \( a^T_j = \mu + \frac{S_j}{2} - \gamma \) with \( j \in \{B, C\} \) and the resulting size of the arbitrage opportunity is given by

\[
Arb_{a^T_j} = b_k - a^T_j = \left( \mu - \frac{S_k}{2} \right) - \left( \mu + \frac{S_j}{2} - \gamma \right) = \frac{S_k + S_j}{2} + \frac{S_k + S_j}{2} + 2c = 2c,
\]

Equivalently, it is the same for bid limit order initiated arbitrage opportunities \( Arb_{b^T_j} = 2c \). Hence, all arbitrage opportunities are expected to have the same size.