Toxic Arbitrage*

Thierry Foucault† Roman Kozhan‡ Wing Wah Tham§

September, 2016

Abstract

Short lived arbitrage opportunities arise when prices adjust with a lag to new information. They are toxic because they expose dealers to the risk of trading at stale quotes. Hence, theory implies that more frequent toxic arbitrage opportunities and a faster arbitrageurs' response to these should impair liquidity. We provide supporting evidence using data on triangular arbitrage. As predicted, illiquidity is higher on days when the fraction of toxic arbitrage opportunities and arbitrageurs' relative speed are higher. Overall, our findings suggest that the price efficiency gain of high frequency arbitrage comes at the cost of increased adverse selection risk.

Keywords: Arbitrage; Liquidity; Adverse Selection; High Frequency Trading.

JEL Classification: D50, F31, G10

*We are grateful to the editor, Andrew Karolyi, and two anonymous referees whose comments helped to improve the paper. We also thank Mark Van Achter, Yacine Aït-Sahalia, Hank Bessembinder, Geir Bjønnes, Michael Brennan, Alain Chaboud, Pierre Collin-Dufresnes, Jean-Edouard Colliard, Matthijs Fleischer, Arie Gozluklu, Carole Gresse, Terry Hendershott, Johan Hombert, Bob Jarrow, Frank de Jong, Pete Kyle, Grace Xing Hu, Olga Lebedeva, Bruce Lehmann, Katya Malinova, Albert Menkveld, Michael Moore, Pamela Moulton, Maureen O'Hara, Marco Pagano, Andreas Park, Joël Peress, Angelo Ranaldo, Vikas Raman, Dagfinn Rime, Fabrice Riva, Gideon Saar, Mark Salmon, Daniel Schmidt, Elvira Sojli, Clara Vega, Kumar Venkataraman, Chen Yao, and Mao Ye. We are also grateful to seminar and conference participants at Bath University, Cornell University, EIEF and Consob, the Norwegian School of Business, Vrije University, Manchester University, KU Leuven, the Autorité des Marchés Financiers, the 2015 American Finance Association meetings, the 8^{th} conference of the Paul Wooley Center for the study of Capital Markets Dysfunctionnality, the workshop on market microstructure theory and applications at Cambridge, the 9^{th} Central Bank Workshop on the Market Microstructure of Financial Markets, the BIRS workshop on modeling high frequency trading activity, the 6^{th} Erasmus Liquidity Conference, the Conference on Liquidity and Arbitrage Trading in Geneva, and the CityU Finance Conference in Hong Kong. Thierry Foucault acknowledges financial support from the Investissements d'Avenir Labex (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

†Corresponding Author: HEC, Paris, 1 rue de la Libération, 78351 Jouy en Josas; tel: +33 139679569; e-mail: foucault@hec.fr

‡Warwick Business School, University of Warwick, Scarman Road, Coventry, CV4 7AL, UK; tel: +44 2476522114; e-mail: Roman.Kozhan@wbs.ac.uk

§School of Banking and Finance, University of New South Wales, UNSW Sydney NSW 2052 Australia; tel: +61 (2)93855484; e.mail: w.tham@unsw.edu.au

Arbitrageurs play a central role in financial markets. When the Law of One Price (LOP) breaks down, they step in, buying the cheap asset and selling the expensive one. Thereby, arbitrageurs enforce the LOP and make markets more price efficient. In theory, arbitrage opportunities should disappear instantaneously. In reality, they do not because arbitrage is not frictionless. As Duffie (2010) points out: *"The arrival of new capital to an investment opportunity can be delayed by fractions of a second in some markets, for example an electronic limit order-book market for equities, or by months in other markets [...]."*

Various frictions (e.g., short-selling costs, funding constraints, idiosyncratic risks, etc.) explain why some arbitrage opportunities persist (see Gromb and Vayanos 2010). For very short-lived arbitrage opportunities–those lasting fractions of a second– attention costs and technological constraints are the main impediments to a seamless Law of One Price. These barriers are falling as high frequency arbitrageurs invest massively to detect and exploit ever faster arbitrage opportunities. Returns on high speed arbitrage are substantial because arbitrage opportunities are very frequent at the time scales of milliseconds (see Budish et al. 2015). This evolution has triggered debates about the social value of high speed arbitrage, and in particular about whether arbitrage strategies "*benefit or harm the interests of long-term investors and market quality [...]*" (U.S. Securities and Exchange Commission (2010), Section B, p.51).

Arbitrageurs can be beneficial or harmful for other investors, depending on the cause of arbitrage opportunities. When these opportunities are due to transient demand or supply shocks ("price pressures"), arbitrageurs implicitly act as liquidity providers in exploiting them (see, for instance, Holden 1995; Gromb and Vayanos 2002, 2010). In this case, trades between arbitrageurs and their counterparties are mutually beneficial.[1] However, short lived arbitrage opportunities are also due to asynchronous adjustments in asset prices following information arrival. Arbitrageurs' profits in these trades are obtained at the expense of dealers with stale quotes.[2] Thus, asynchronous price adjustments to information in asset pairs generate "toxic" arbitrage opportunities, i.e., opportunities in which dealers are at risk of being adversely selected.[3] High speed arbitrageurs can harm market liquidity through this channel because dealers

[1] For instance, Gromb and Vayanos (2002) write (on p.362): *"In our model, arbitrage activity benefits all investors. This is because through their trading, arbitrageurs bring prices closer to fundamentals and supply liquidity to the market."*

[2] This problem is not new. For instance, in the 90s, professional day traders (so-called SOES bandits) were picking off Nasdaq dealers with stale quotes by using Nasdaq's Small Order Execution System (a system that guaranteed automatic execution of market orders up to a certain size at Nasdaq dealers quotes). See Harris and Schultz (1997) and Foucault et al. (2003).

[3] Our definition of a toxic trade follows Easley et al. (2012). They write (p.1458): *"Order flow is regarded as toxic when it adversely selects market makers who may be unaware that they are providing liquidity at a loss."*

charge larger bid-ask spreads to cover the risk of trading at stale quotes (Copeland and Galai 1983).

Our contribution is to model this channel and provide evidence of its importance for liquidity. To our knowledge, our paper is first to do so. This is important for at least two reasons. First, arbitrage is a central notion in finance. Thus, understanding how it affects market quality in general, not just pricing efficiency, is of broad interest. Second, recent proposals advocate slowing down the pace of trading precisely on the grounds that high speed arbitrageurs raise dealers' risk of trading at stale quotes (see, e.g., Budish et al. 2015). However, there is yet no evidence on whether arbitrageurs' contribution to this risk is significant or not. Measuring this contribution is not straightforward because it is not the level of arbitrage activity per se that should affect dealers' risk of trading at stale quotes (and therefore their spreads). Rather, as shown by our model, this risk is determined both by the "*arbitrage mix*" (i.e., the proportion of toxic arbitrage opportunities in the pool of all arbitrage opportunities) and arbitrageurs' *relative speed* of reaction to toxic arbitrage opportunities. Specifically, illiquidity should be higher in periods (or for asset pairs) where (i) the fraction of arbitrage opportunities that are toxic is higher (the arbitrage mix is more toxic) or (ii) the likelihood that a toxic arbitrage opportunity terminates with an arbitrageur's trade is higher (arbitrageurs are relatively faster).

These two predictions follow from a new model of cross-market arbitrage that we develop in the first part of our paper. In the model, arbitrage opportunities can be either toxic (due to asynchronous price adjustments to news) or non-toxic (due to liquidity shocks). As in reality, an arbitrage opportunity terminates either with an arbitrageur's trade or a dealer's quote update, depending on whoever observes the opportunity first. We solve for equilibrium bid-ask spreads in each asset and traders' optimal speed of reaction to arbitrage opportunities. Thus, in equilibrium, illiquidity and the duration of arbitrage opportunities (a measure of pricing efficiency) are jointly determined.

The model generates predictions (i) and (ii) above and two additional predictions about the durations of arbitrage opportunities. First, when the arbitrage mix becomes more toxic, arbitrage opportunities should be shorter, even though bid-ask spread costs of arbitrage are higher. The reason is that dealers react faster to arbitrage opportunities (by updating their quotes) when they expect more of them to be toxic. This effect induces arbitrageurs to be faster as well and, as a result, arbitrage opportunities are more short-lived. Second, by a similar logic, a technological change that makes arbitrageurs relatively faster should reduce the duration of arbitrage opportunities, even though it increases illiquidity.

We test these predictions using data on triangular arbitrage opportunities for three currency pairs (dollar-euro, dollar-pound, and pound-euro).[4] Although our predictions and methodology apply to any type of high frequency arbitrage opportunities, we focus on triangular arbitrage opportunities for a couple of reasons.

The first one is practical. For our tests, we must accurately measure when an arbitrage begins, when it terminates, how it terminates (with a trade or a quote update), and we must track prices after the arbitrage terminates (to identify toxic arbitrage opportunities; see below). Our data has the required granularity for this analysis: we observe all orders and trades for currency pairs in our sample from January 2003 to December 2004 in Reuters D-3000 (one of the two major interdealer trading platforms used by foreign exchange dealing banks) with a time stamp accuracy of 10 milliseconds.[5] Moreover, asynchronicities in price reporting for different assets is not an issue in our data because all data are generated by the same trading platform.

Second, strategies exploiting triangular arbitrage opportunities are not hindered by taxes, short-selling or funding constraints, and the risk of these strategies is very limited. Hence, standard limits to arbitrage cannot explain why triangular arbitrage opportunities are not eliminated instantaneously (see Pasquariello 2014).[6] The most likely explanation is that, as in our model, technological constraints limit the speed at which traders react to arbitrage opportunities. Thus, triangular arbitrage opportunities are very similar to other high speed opportunities: they are (i) frequent (we observe more than 172,044 in our sample), (ii) very short-lived (in our sample, 25% of all arbitrage opportunities last less than half a second), (iii) more efficiently exploited by machines than by humans, and (iv) they deliver very thin profits per opportunity (0.6 to 0.7 basis points in our sample).[7]

As any other arbitrage opportunities, triangular arbitrage opportunities arise for two reasons: (i) asynchronous price adjustments of different currency pairs to new information or (ii)

[4] One can buy euros with dollars, exchange the euros against pounds, and then exchange pounds against dollars. If one ends up with more dollars than the initial dollar investment then a triangular arbitrage opportunity exists. We define triangular arbitrage opportunities formally in Section 2.2.

[5] Kozhan and Tham (2012) use the same data to measure the profitability of triangular arbitrage opportunities.

[6] Pasquariello (2014) finds that none of the usual proxies for limits to arbitrage explain the size of triangular arbitrage opportunities (Pasquariello 2014, see Table 2 in). For other relatively short-lived opportunities, these limits can be more important. For instance, Gagnon and Karolyi (2010) find that holding costs (e.g., idiosyncratic risk) explain the size of arbitrage opportunities between home and U.S. stock prices for stocks cross-listed in the U.S.

[7] Arbitrage opportunities in the foreign exchange market (either violations of covered interest parity or triangular arbitrage) are well documented. See, for instance, Akram et al. (2008), Fong et al. (2008), Fenn et al. (2009), Mancini-Griffoli and Ranaldo (2011), Marshall et al. (2008), Kozhan and Tham (2012), Ito et al. (2013), Chaboud et al. (2014), and Pasquariello (2014).

price pressures effects. Price pressure effects generate price reversals, while asynchronous price adjustments to information generate staggered price movements in the same direction for related assets.[8] Thus, as in Schultz and Shive (2010), we use price patterns following the occurrence of arbitrage opportunities to sort them into two groups: toxic (characterized by staggered price movements in the same direction following the occurrence of a triangular arbitrage opportunity) and non-toxic (characterized by a reversal in the rate of the currency pair that triggers the arbitrage opportunity).[9] With this approach, we obtain 83,488 toxic arbitrage opportunities (about 112 per day), i.e., about 48% of all arbitrage opportunities in the sample. On average (across all days in our sample), we find that these opportunities terminate with an arbitrageurs' trade in about two-third of the cases. Thus, arbitrageurs are often faster than dealers in our sample.

Arbitrageurs' relative speed is endogenous to illiquidity because arbitrageurs have less incentive to quickly detect arbitrage opportunities when transactions costs are high. To account for this in our tests, we use an instrument for arbitrageurs' relative speed (measured by the frequency with which a toxic arbitrage opportunity terminates with a trade). Until July 2003, traders had to manually submit their orders to Reuters D-3000. In July 2003, Reuters introduced the "AutoQuote API" functionality (API means "Application Programming Interface"). Traders using this functionality can directly feed their algorithms to Reuters D-3000 data and let these submit orders automatically, reducing thereby their monitoring costs. Arbitrageurs were among the first to use Autoquote API (see Chaboud et al. 2014), which suggests that the reduction in monitoring costs mainly accrued to them.[10] In our model, this implies that arbitrageurs' relative speed should increase following the introduction of Autoquote API. Thus, we instrument arbitrageurs' relative speed with AutoQuote API. In line with our conjecture, the

[8] For instance, for cross-listed stocks, Gagnon and Karolyi (2009) show that there is negative autocorrelation in home and foreign returns at the daily frequency. This negative autocorrelation however is weaker for stocks in which informed trading is more intense. This is consistent with the idea that delays in adjustment to information for assets with correlated payoffs (e.g., cross listed stocks) induce positive spillovers in price changes.

[9] Suppose that euro/dollar dealers receive information that calls for an appreciation of the euro and raise their bid and ask quotes (expressed in dollars per euro). If this appreciation is large enough and dealers in, say, the dollar/pound market are slow to adjust their quotes, there is a triangular arbitrage opportunity: one can indirectly buy dollars with euros at a price less than the current bid price in the dollar/euro market. This toxic arbitrage opportunity vanishes when dealers in the dollar/pound market raise their quotes or arbitrageurs hit stale quotes in this market. In either case, the rate in the dollar/pound market adjusts in the direction of the shift in the euro/dollar market. Alternatively, if euro/dollar dealers temporarily accumulate a large short position in euro, they will mark up the value of the euro against the dollar to attract sellers of euros and reduce their risk exposure. If this price pressure effect is large enough, a non-toxic triangular arbitrage opportunity arises. As dealers' short position decreases, their quotes will revert (see, for instance, Grossman and Miller 1988).

[10] Hendershott et al. (2011) use the implementation of the NYSE "autoquote" software in 2003 as an instrument for algorithmic trading. The NYSE autoquote functionality is different from Reuters AutoQuote API because the former automates the dissemination of updates in best quotes for NYSE stocks while the latter automates order entry. Automation of order entry clearly accelerates the speed at which traders react to market events. We discuss the differences between our findings and those in Hendershott et al. (2011) in Section 3.1.

first stage of the IV regression shows that the introduction of Autoquote API has a significant positive effect on the likelihood that a toxic arbitrage opportunity terminates with a trade.

More importantly, the second stage shows that, as predicted, the likelihood that a toxic arbitrage opportunity terminates with a trade has a positive effect on illiquidity. For instance, a 1% increase in this likelihood in a day is associated with a 0.063 basis points increase in quoted bid-ask spreads in this day (2.3 to 5% of the average bid-ask spread depending on the currency pair). The economic size of this effect is significant given the daily trading volume for the currency pairs in our sample (we estimate that a 0.063 basis points increase in quoted spread raises the total cost of trading for the currency pairs in our sample by about $\$131,319$ per day). We find similar effects when we measure illiquidity with effective spreads, the slope of limit order books, or a measure of adverse selection costs for dealers.

Moreover, consistent with our first prediction, we also find a positive and significant relation between the daily fraction of arbitrage opportunities that are toxic and illiquidity. Specifically, on days where this fraction is higher, illiquidity is higher, after controlling for the number of arbitrage opportunities (scaled by the number of trades) and standard determinants of illiquidity. For instance, a one standard deviation increase in the fraction of arbitrage opportunities that are toxic on one day is associated with a 2.3% increase in the average quoted spread for the currencies in our sample on the same day. Thus, the arbitrage mix matters: illiquidity is higher when arbitrage opportunities are more frequently due to asynchronous price adjustments than price pressures.

In sum, consistent with our predictions, illiquidity is positively related to (i) the fraction of arbitrage opportunities that are toxic and (ii) arbitrageurs' relative speed. Additional predictions of our model are supported by the data as well: (a) the duration of arbitrage opportunities is shorter on days in which the fraction of arbitrage opportunities that are toxic is higher and (b) the introduction of AutoQuote API (an increase in arbitrageurs' relative speed) coincides with a 6.7% (about 115 milliseconds) decrease in the average duration of arbitrage opportunities in our sample.

It is well known that liquidity facilitates arbitrage. The reverse relation–the effect of arbitrageurs on liquidity (our focus here)–has received much less attention.[11] Kumar and Seppi

[11] Roll et al. (2007) show that there exist two-way relations between index futures basis and stock market liquidity. In particular, a greater index futures basis Granger-causes greater stock market illiquidity. Roll et al. (2007) argue that this effect could be due to arbitrageurs's trades but do not specifically show that these trades explain the relation. Rosch (2014) uses the size of arbitrage opportunities in Depositary Receipts as an inverse proxy for arbitrage activity and finds a positive association between arbitrage activity and liquidity.

(1994) model cross-market arbitrageurs as informed traders and show that, in their model, the effect of the number of arbitrageurs on liquidity is non-monotonic. They do not study how the arbitrage mix affects illiquidity (all arbitrage opportunities are due to stale quotes in their model) and traders' speed is not a choice variable in their model. Hence, they do not derive the predictions that we test in this paper.

Several papers argue that fast traders raise adverse selection costs for slow traders.[12] Our empirical findings about the effect of arbitrageurs' speed are consistent with this view. The main message of our paper, however, is not that adverse selection is a source of illiquidity. This, of course, is well known. What is novel is that high speed arbitrage can be a source of adverse selection and that, for this reason, the "arbitrage mix" in an asset pair is a determinant of its liquidity. These new findings contribute to the burgeoning literature on short lived arbitrage opportunities in various asset pairs, such as currencies, ETFs, cross-listed stocks, dual class shares etc. (see, for instance, Akram et al. 2008; Ben-David et al. 2012; Gagnon and Karolyi 2010; or Schultz and Shive 2010).

1. Illiquidity, arbitrage mix, and arbitrageurs' relative speed

In this section, we present the model of cross-market arbitrage that guides our empirical analysis. Before describing it formally, it is worth outlining its main ingredients and why these are required for our analysis. The model has two assets with identical payoffs. Quotes for each asset are posted by two different market makers. To generate arbitrage opportunities, we assume that the market maker in one asset can receive a random shock to his valuation for this asset, either due to news arrival or liquidity needs. This feature enables us to study how the likelihood that an arbitrage is toxic (i.e., due to news arrival) affects market makers' bid-ask spreads (illiquidity). Moreover, we allow for heterogeneity in traders' speeds of reaction to arbitrage opportunities. This feature is required for analyzing how arbitrageurs' relative speed affects illiquidity. Importantly, we endogenize traders' speeds because, in reality, their incentive to eliminate arbitrage opportunities quickly is endogenous to illiquidity. Finally, when there is no shock to market makers' valuation, we assume that market makers trade with liquidity traders to capture the fact that, in practice, arbitrageurs only account for a fraction of trading volume.

[12] See, for instance, Biais et al. (2015), Brogaard et al. (2015) Budish et al. (2015), Foucault et al. (2003), Foucault et al. (2016), Garvey and Wu (2010), Hendershott and Moulton (2011), Hoffmann (2014), Jovanovic and Menkveld (2012), or Menkveld and Zoican (2014).

1.1 Model

The model has two assets, X and Y, three dates ($t \in \{0, 1, 2\}$), two market makers, and one arbitrageur (A). At date $t = 2$, the payoffs of the assets, denoted $\tilde{\theta}_X$ for X and $\tilde{\theta}_Y$ for Y are realized. These payoffs are identical and given by $\tilde{\theta}_Y = \tilde{\theta}_X = \mu + \tilde{\varepsilon}$ where $\tilde{\varepsilon} = \sigma/2$ or $\tilde{\varepsilon} = -\sigma/2$ with equal probabilities, where $\sigma > 0$.[13]

Market Makers. Market maker $j \in \{X, Y\}$ is specialized in asset j. As in other models of multi-asset trading (e.g., Boulatov et al. 2013 or Pasquariello 2016), markets for assets X and Y are segmented because market makers in each asset are different and, for this reason, information available to one market maker (e.g., from news or past trades) is not instantaneously available to the other.[14] Thus, short lived arbitrage opportunities between markets X and Y can happen. In the baseline version of the model, we focus on the case in which shocks to market maker Y's valuation (due to information arrival or liquidity needs) cause these opportunities (see below). Thus, asset Y "leads" asset X.

At date $t = 1$, market makers simultaneously post an ask price, a_j, and a bid price, b_j, for $j \in \{X, Y\}$ such that:

$$a_j = v_j + \frac{S_j}{2}, \quad \text{and} \quad b_j = v_j - \frac{S_j}{2}, \tag{1}$$

where v_j is market maker j's valuation for asset j and S_j is the bid-ask spread for asset j. Quotes are for a fixed number of shares, normalized to 1, of each asset.

Market makers' valuations for assets X and Y are determined at date 0. Market maker X derives a utility $\tilde{\theta}_X$ per share of asset X owned at date 2 and has no information about the payoff of asset X. Thus, prior to trading, his valuation is $v_X = \mathsf{E}(\tilde{\theta}_X) = \mu$.

With probability $(1 - \alpha)$, market maker Y also derives a utility $\tilde{\theta}_Y$ per share of asset Y owned at date 2 and has no information about the payoff of asset Y. In this case, his valuation is $v_Y = \mathsf{E}(\tilde{\theta}_Y) = \mu$. Alternatively, with probability α, there is a shock to market maker Y's valuation. This shock can be due either to information arrival (with probability φ) or liquidity needs (with probability $(1 - \varphi)$). In case of information arrival, market maker Y privately observes ε and therefore his valuation for asset Y becomes $v_Y = \mathsf{E}(\tilde{\theta}_Y \mid \varepsilon) = \mu + \varepsilon$. In case of a

[13] For instance, assets X and Y might be two derivatives on the same underlying asset (e.g., the E-mini S&P 500 Futures (ES) and the SPDR S&P 500 Exchange Traded Funds (SPY)) or asset X might be a synthetic asset with the same payoff as asset Y.

[14] In reality, information flows between markets cannot be instantaneous and market makers are specialized. For instance, market makers in equities markets specialize in a few individual stocks and do not share information in real time, even when they belong to the same trading desk (see Naik and Yadav 2003). This is also the case in currency markets where market makers often specialize in one currency pair (see Bjønnes and Rime 2005).

liquidity need, market maker Y's utility for the asset becomes $(\tilde{\theta}_Y + \tilde{\delta})$ per share of asset Y owned at date 2 where $\tilde{\delta}$ is equal to $\sigma/2$ or $-\sigma/2$ with equal probabilities and is independent from $\tilde{\varepsilon}$. Thus, in this case, market maker Y's valuation for the asset becomes $v_Y = \mathsf{E}(\tilde{\theta}_Y) + \delta = \mu + \delta$. The private value component δ represents, for instance, the hedging value of the asset for the market maker (as in Duffie et al. 2005).

Arbitrage Opportunities. As assets X and Y have identical payoffs, the arbitrageur can take advantage of a divergence in market makers' valuations. For instance, suppose that $v_Y > v_X$. If the arbitrageur buys asset X (at ask price a_X) and sells asset Y (at bid price b_Y), she locks in a sure profit of $b_Y - a_X = (v_Y - v_X) - (S_Y + S_X)/2$ since assets X and Y have identical payoffs. By symmetry, if $v_X > v_Y$, the arbitrageur's profit is $(v_X - v_Y) - (S_Y + S_X)/2$. Thus, if she trades, the arbitrageur's profit is:

$$ArbProfit = \Delta_{XY} - (S_Y + S_X)/2, \tag{2}$$

where $\Delta_{XY} = |v_Y - v_X|$. The arbitrageur's profit is positive if $\Delta_{XY} > (S_Y + S_X)/2$, i.e., if the difference in market makers' valuations is large enough relative to the bid-ask spread cost borne by the arbitrageur.

Given our assumptions, $\Delta_{XY} = \sigma/2$ when there is a shock to market maker Y's valuation, whether this shock is due to news arrival or a liquidity need.[15] Thus, in these cases, the arbitrageur's expected profit (eq.(2)) is strictly positive when

$$S_X + S_Y < \sigma. \tag{3}$$

This condition will always be satisfied in equilibrium (see below).

If there is no shock to market maker Y's valuation (probability $(1-\alpha)$) then $\Delta_{XY} = 0$ and there is no profitable trade for the arbitrageur. In this case, a liquidity trader arrives in the market to buy or sell one share of asset X or Y, with equal probabilities.

Toxic and non-toxic arbitrage opportunities. If market maker X trades with a liquidity trader, he earns half his bid-ask spread. In contrast, if he trades with the arbitrageur, his expected profit depends on the type of shock that triggers the arbitrage opportunity.

[15] For instance, if this shock is positive, we have $v_Y = \mu + \sigma/2$. Accordingly, in this case, $\Delta_{XY} = (v_Y - v_X) = \mu + \sigma/2 - \mu = \sigma/2$. The absolute difference in market makers' valuation is the same whether news arrival or liquidity needs trigger the change in market maker Y's valuation because innovations in the asset value are identical (equal to σ) in each case. This assumption can be relaxed but it simplifies the exposition of the model by reducing the number of parameters.

For instance, consider a positive shock to market maker Y's valuation. In this case, if the arbitrageur trades, he buys asset X. If the shock to Y's valuation is due to news then asset X's expected payoff is $\mathsf{E}(\tilde{\theta}_X \mid \tilde{\varepsilon} = \sigma/2) = \mu + \sigma/2$. Thus, if market maker X sells the asset to the arbitrageur, he earns an expected profit of $a_X - (\mu + \sigma/2) = (S_X - \sigma)/2$, which is negative if the arbitrageur finds it profitable to trade (i.e., if Condition (3) is satisfied). In other words, an arbitrage opportunity due to the arrival of news about asset Y is toxic for market maker X because he is exposed to the risk of trading at a loss with the arbitrageur.

If instead, the shock to market maker Y's valuation is due to a liquidity shock then the expected payoff of asset X is $\mathsf{E}(\tilde{\theta}_X \mid \tilde{\delta} = \sigma/2) = \mu$. Thus, if market maker X trades with the arbitrageur, he obtains an expected profit of $a_X - \mu = \frac{S_X}{2}$, as if he were trading with a liquidity trader. Thus, an arbitrage opportunity due to a liquidity shock for market maker Y is non-toxic for market maker X.

Market maker X can avoid toxic trades if he cancels his quotes before the arbitrageur hits them. We denote by $(1 - \pi)$ the probability that X is fast enough to do so. Thus, when a toxic arbitrage opportunity occurs, the arbitrageur can actually trade on it with probability π. In contrast, if a non-toxic arbitrage opportunity happens, there is no reason for market maker X to cancel his quotes since he makes a profit when he trades with the arbitrageur. Thus, we assume that the arbitrageur can exploit a non toxic arbitrage opportunity with certainty.

Table 1 gives the expected payoffs of the arbitrageur and each market maker for each possible event at date 1: a liquidity trader arrives (probability $(1 - \alpha)$); a toxic arbitrage happens (probability $\alpha\varphi$); a non-toxic arbitrage happens (probability $\alpha(1 - \varphi)$). Thus, parameter α controls the likelihood of occurrence of an arbitrage opportunity while parameter φ controls the likelihood that an arbitrage is toxic conditional on an arbitrage opportunity occurring. We therefore call it the "arbitrage mix".

Market maker Y's quotes always reflect all available information about the payoff of asset Y. Thus, in contrast to market maker X, he earns half his bid-ask spread in all cases. When a toxic arbitrage opportunity happens, total gains from trade between the arbitrageur and market makers (last line of Table 1) are zero: the loss of market maker X if he trades with the arbitrageur is just equal to the profits of the arbitrageur and market maker Y. In contrast, when a non-toxic arbitrage opportunity happens, total gains from trade are strictly positive and equal to $\sigma/2$, the difference between market makers' valuations. Indeed, in this case, this difference reflects true gains from trade between market makers. By trading across the two markets, the arbitrageur enables market makers to achieve these gains.

Table 1: Traders' Expected Payoffs

This table gives the expected payoff of each trader in the model for the various possible events at date 1. The line called "Termination" indicates how the trading round terminates: a trade (from a liquidity trader or the arbitrageur) or a cancellation of his quotes by market maker X. We use the following abbreviations: "Liq. trader" for "Liquidity trader"; "Arb." for "Arbitrageur"; and "prob" for probability. The last line of the table gives the sum of the expected payoffs for the arbitrageur and the market makers for each event considered in the table.

	Liq. trader arrives (with prob $1-\alpha$)	A toxic arbitrage happens (with prob $\alpha\varphi$)		A non-toxic arbitrage happens (with prob $\alpha(1-\varphi)$)	
	$\downarrow$	$\swarrow$	$\searrow$	$\swarrow$	$\searrow$
Termination	Liq. trader trades	Arb. trades (with prob π)	X cancels (with prob $1-\pi$)	Arb. trades (with prob 1)	X cancels (with prob 0)
Arb's expected payoff	0	$\frac{\sigma-(S_X+S_Y)}{2}$	0	$\frac{\sigma-(S_X+S_Y)}{2}$	0
X's expected payoff	$\frac{S_X}{4}$	$-\frac{(\sigma-S_X)}{2}$	0	$\frac{S_X}{2}$	0
Y's expected payoff	$\frac{S_Y}{4}$	$\frac{S_Y}{2}$	0	$\frac{S_Y}{2}$	0
Aggregate expected payoff (Arbs + Market Makers)	$\frac{S_X+S_Y}{4}$	0	0	$\frac{\sigma}{2}$	0

Speed. In reality, the probability, π, that an arbitrageur can hit stale quotes before they are cancelled depends on her speed of reaction to market events relative to liquidity providers. Speed is a choice variable for traders. We model it as in Foucault et al. (2003). Specifically, when an arbitrage opportunity occurs, it takes D^a and D^m units of time for the arbitrageur and market maker X, respectively, to spot it (index m refers to a market maker), where D^a and D^m are exponentially distributed with intensity γ and λ, respectively. The randomness of D^a and D^m captures the fact that, in practice, traders' response times to market events (e.g., an arbitrage opportunity) depend on a myriad of random factors (e.g., time required by platforms to process orders) that cannot be fully controlled by traders.

If $D^a < D^m$, the arbitrageur is first to observe the arbitrage opportunity and she exploits it. Otherwise market maker X is first to observe the opportunity and cancels his quotes. Thus, the likelihood, π, that a toxic arbitrage opportunity terminates with a trade by the arbitrageur is:

$$\pi = \Pr\left(D^a < D^m\right) = \frac{\gamma}{\lambda+\gamma}, \tag{4}$$

where the second equality follow from the fact that D^m and D^a are exponentially distributed. Other things equal, an increase in γ (resp., λ) increases (resp., reduces) the likelihood that the arbitrageur exploits a toxic arbitrage before it vanishes. We therefore refer to λ and γ as traders' speeds. In line with intuition, π increases with the arbitrageur's *relative* speed, i.e., $\frac{\gamma}{\lambda}$.

Technological investments (e.g., in hardware and software, fast data feed, dedicated communication lines, colocation, etc.) and attention enable traders to reduce their average response times to market events (e.g., they can reduce latencies in communicating with trading platforms). Investments and attention are costly.[16] Hence, we assume that if the market maker operates at speed λ then he bears a monitoring cost $C^m(\lambda) = \frac{c^m\lambda}{2}$. Similarly, if the arbitrageur operates at speed γ then she bears a cost $C^a(\gamma) = \frac{c^a\gamma}{2}$.[17]

Traders' expected profits. Using Table 1, we deduce that the expected profits of market makers X, Y, and the arbitrageur, net of monitoring costs, are:

$$\Pi^X(S_X, \lambda, \gamma) = -(\varphi\alpha\pi)\frac{(\sigma - S_X)}{2} + (1 - \alpha(2\varphi - 1))\frac{S_X}{4} - \frac{c^m\lambda}{2}, \tag{5}$$

$$\Pi^Y(S_Y, \lambda, \gamma) = (2\alpha(1 - (1 - \pi)\varphi) + (1 - \alpha))\frac{S_Y}{4}, \tag{6}$$

$$\Pi^A(S_X, S_Y, \lambda, \gamma) = (\alpha\varphi\pi)\frac{(\sigma - (S_X + S_Y))}{2} + \alpha(1 - \varphi)\frac{(\sigma - (S_X + S_Y))}{2} - \frac{c^a\gamma}{2}, \tag{7}$$

where, as previously explained, $\pi = \frac{\gamma}{\lambda+\gamma}$.[18] For instance, the first term in eq.(5) is market maker X's expected losses when he trades with the arbitrageur in a toxic arbitrage opportunity while the second term is his expected payoff when he trades with liquidity traders or the arbitrageur in a non-toxic arbitrage. Finally, the last term is the cost of speed for the market maker.

Equilibrium. Market makers and the arbitrageurs simultaneously choose their bid-ask spreads and speeds. We focus on competitive equilibria, i.e., a set $\{S_X^*, S_Y^*, \lambda^*, \gamma^*\}$ such that:

$$\Pi^Y(S_Y^*, \lambda^*, \gamma^*) = 0, \tag{8}$$

$$\Pi^X(S_X^*, \lambda^*, \gamma^*) = 0, \tag{9}$$

$$\Pi^X(S_X, \lambda, \gamma^*) < 0 \quad \forall S_X < S_X^*, \forall\lambda, \tag{10}$$

$$\lambda^* \in Argmax_\lambda \quad \Pi^X(S_X^*, \lambda, \gamma^*), \tag{11}$$

[16] Attention can be interpreted literally as the effort that human traders must exert to follow prices in different markets. It can also represent the computing capacity that traders allocate to a particular task, e.g., detecting an arbitrage opportunity in a specific pair of assets. Allocating greater capacity to this specific pair reduces the capacity available for other trading opportunities, which generates an opportunity cost.

[17] We assume linear costs of speed to obtain closed form solutions. Implications however do not crucially depend on this assumption. For instance, predictions of the model are identical with quadratic costs of speeds.

[18] Other things equal, traders' expected profits gross of their investment in speed increase in their relative speed (i.e., arbitrageurs' expected profits increase in π while dealers' expected profits increase with π^{-1}). This is consistent with Baron et al. (2016), who find that high frequency trading firms' revenues increase in their relative speed, not absolute speed.

$$\gamma^* \in Argmax_\gamma \quad \Pi^A(S_X^*, S_Y^*, \lambda^*, \gamma). \tag{12}$$

In a competitive equilibrium, market makers earn a zero expected profit and their bid-ask spread cannot be profitably undercut. For market maker Y, these two conditions are satisfied if the zero profit condition (8) holds because market maker Y's expected profit increases in his bid-ask spread. For market maker X, the zero profit condition (9) is necessary but not sufficient. In addition, his bid-ask must be such that it cannot be profitably undercut by choosing a speed level different from his equilibrium speed level (Condition (10)). Last, market maker X and the arbitrageur choose the speed that maximizes their expected profit given other traders' equilibrium actions (Conditions (11) and (12)).

Proposition 1 provides a closed form solution for the unique competitive equilibrium of the model. It forms the backbone of our empirical tests (see the next section). Henceforth, $\rho = \left(\frac{c^m}{c^a}\right)$ denotes the ratio of market maker X's cost of speed to the arbitrageur's cost of speed.

Proposition 1. *There is a unique competitive equilibrium. In this equilibrium, market maker Y's bid-ask spread is nil ($S_Y^* = 0$) while market maker X's bid-ask spread is:*

$$S_X^* = \left(\frac{2\varphi\alpha\pi^*(\rho)(2-\pi^*(\rho))}{2\varphi\alpha\pi^*(\rho)(2-\pi^*(\rho)) + (1-\alpha(2\varphi-1))}\right) \times \sigma. \tag{13}$$

where the likelihood that a toxic arbitrage opportunity terminates with an arbitrageur's trade is:

$$\pi^*(\rho) = \frac{\gamma^*}{\lambda^* + \gamma^*} = \frac{\rho}{1+\rho}. \tag{14}$$

Moreover, in equilibrium, absolute speeds for market maker X and the arbitrageur are:

$$\lambda^* = \frac{\varphi\alpha(\sigma - S_X^*)c^a}{(c^a + c^m)^2}, \tag{15}$$

$$\gamma^* = \frac{\varphi\alpha(\sigma - S_X^*)c^m}{(c^a + c^m)^2}. \tag{16}$$

Market maker Y's competitive bid-ask spread is zero because he is never exposed to the risk of trading at stale quotes. In contrast, market maker X's competitive bid-ask spread is strictly positive, except if he is not exposed to the risk of trading at stale quotes (e.g., if $\varphi = 0$ or $\alpha = 0$). His bid-ask spread allows him to recoup his losses when he trades at stale quotes with gains in other cases. For $\varphi < 1$, market maker X's bid-ask spread is always strictly smaller than σ. Thus, as mentioned previously, in equilibrium, total bid-ask spread costs for the arbitrageur

$(S_X^* + S_Y^*)$ are smaller than σ (i.e., Condition (3) is satisfied in equilibrium). Interestingly, this is the case even when $\alpha = 1$, i.e., when trades happen only between the arbitrageur and market makers (no liquidity traders). The reason is that strictly positive gains from trade exist when a non-toxic arbitrage opportunity occurs. Thus, in this case, all parties can trade at a profit, which allows market maker X to recoup his trading losses when toxic arbitrage opportunities occur. The more general case in which $\alpha < 1$ allows us to capture the fact that arbitrageurs' trades only account for a fraction of the trading volume in reality.

In equilibrium, the likelihood that a toxic arbitrage terminates with an arbitrageur's trade, π^*, increases with ρ, the ratio of market maker X's cost of speed (c^m) to the arbitrageur's cost of speed (c^a). The arbitrageur is relatively faster than the market maker ($\pi^* > 1/2$) if and only if her cost of speed is relatively smaller ($\rho > 1$). Our predictions do not depend on who is relatively faster (i.e., do not specifically require $\rho > 1$).

1.2 Testable predictions

Proposition 1 has several testable implications. First, market maker's X's bid-ask spread (eq.(13)) increases when arbitrage opportunities are more likely to be toxic (φ is higher) or, holding this likelihood constant, when the arbitrageur becomes relatively faster, i.e., when π^* is higher. Traders' speeds however are endogenous and jointly determined with the bid-ask spread (see eq.(15) and (16)). The model suggests using shocks to the arbitrageur's relative cost of speed as a source of exogenous variations for π^*. Indeed, a decrease in the relative cost of speed for the arbitrageur (i.e., an increase in $\rho = c^m/c^a$) triggers an increase in π^* (eq.(14)) and, through this channel only, an increase in the bid-ask spread. These observations yield the following testable implications.

Implication 1. *Consider a pair of assets X and Y linked by a no-arbitrage relation. An increase in the likelihood that an arbitrage opportunity is toxic (φ) causes an increase in the bid-ask spread of asset X.*

Implication 2. *Consider a pair of assets X and Y linked by a no-arbitrage relation. A reduction in arbitrageurs' relative cost of speed (i.e., an increase in ρ) triggers an increase in π^* – the probability of an arbitrageur's trade, conditional on the occurrence of a toxic arbitrage – and, through this channel, it increases the bid-ask spread of asset X.*

In practice, changes in trading technologies can affect traders' relative costs of speed. For instance, the implementation of its "Hybrid Market" mechanism by the NYSE (in 2006) raised off-floor traders' relative speed advantage by increasing twofold the speed of execution of their market orders (see Figure 2 in Hendershott and Moulton (2011)). Another example are speed bumps that delay the execution of incoming market orders, as implemented recently by some trading platforms (e.g., EBS in FX markets or IEX in equities markets).[19] This practice reduces liquidity takers' speed advantage and is thus similar to reducing ρ in our model. In sum, changes in trading technologies that affect market makers' relative costs of speeds can provide good instruments to measure the effect of π^* on the bid-ask spread because they should affect liquidity only through their effects on π^*, as implied by Implication 2. As explained in Section 3.1, we use this insight in our tests.

In our model, the expected duration of an arbitrage opportunity is *jointly* determined with illiquidity (because traders' speeds are inversely related to the equilibrium bid-ask spread; see eq.(15) and (16)) and affected by the same exogenous factors. Thus, in our tests, we will also study how the likelihood that an arbitrage is toxic (φ) and arbitrageurs' relative cost advantage (ρ) affect the expected duration of an arbitrage opportunity, denoted $\mathsf{E}(D)$. According to the model:

$$\mathsf{E}(D) = \varphi\,\mathsf{E}\left(\mathrm{Min}\{D^a, D^m\}\right) + (1-\varphi)\,\mathsf{E}(D^a) = \frac{(1+\rho)-\varphi}{(\gamma^*+\lambda^*)\rho}, \tag{17}$$

where the second equality follows from the fact that D^a and D^m are exponentially distributed with parameters γ^* and λ^*, respectively. As speed is costly, equilibrium speeds are never infinite. Thus, in equilibrium, arbitrage opportunities do not immediately vanish, i.e., $\mathsf{E}(D) > 0$. However, holding arbitrageurs' relative cost advantage, ρ, constant, arbitrage opportunities become increasingly short-lived when c^a and c^m tend to zero because then the sum of traders' absolute speed levels $(\gamma^* + \lambda^*)$ becomes increasingly large.[20] Thus, arbitrage opportunities can be very short-lived in equilibrium, as we will observe in our data.

Now, consider a change in trading technology that reduces the arbitrageur's cost of speed, c^a and thereby increases the arbitrageur's relative cost advantage, ρ. Holding ρ fixed, the direct effect of this change is to induce the arbitrageur to be faster, which raises traders' aggregate speed $(\lambda^* + \gamma^*)$ and therefore reduces the duration of arbitrage opportunities. However, as ρ increases as well, the equilibrium bid-ask spread is higher (Implication 2), which reduces the arbitrageur's expected profit per opportunity and thereby her incentive to be fast (see eq.(14)).

[19]See "*EBS takes new steps to rein in high frequency traders*," Reuters, August 23, 2013 or "*Brad Katsuyama's next chapter*," Bloomberg Markets magazine, October 2015.

[20]Indeed, $\lambda^* + \gamma^* = \frac{\varphi\alpha(\sigma - S_X^*)}{(c^a+c^m)^2}$ and therefore becomes infinite as c^m and c^a go to zero, holding ρ (hence S_X^*) constant

This indirect effect tends to increase the arbitrage duration. However we show in the appendix that it is always dominated by the direct effect. This yields our next testable implication.

Implication 3. *The average duration of arbitrage opportunities should decrease following a decrease in arbitrageurs' cost of speed, c^a (see the appendix for a proof).*

Chaboud et al. (2014) find that algorithmic trading leads to fewer triangular arbitrage opportunities per second in the FX market. They also find that this reduction is mainly due to algorithmic arbitrageurs hitting quotes of slower (human) traders. These findings are consistent with our model if algorithmic trading reduces relatively more the cost of being fast for arbitrageurs. Indeed, in this case, algorithmic trading is associated with an increase in ρ, and therefore π^*, increases. Accordingly, arbitrage opportunities terminate more frequently with arbitrageurs hitting quotes, as found by Chaboud et al. (2014) and, per Implication 3, the duration of arbitrage opportunities decreases. Hence, if one checks for the presence of arbitrage opportunities at fixed points in time (e.g., every second), fewer triangular arbitrage opportunities are observed (even though the true occurrence rate of these opportunities, α, is unchanged).

Implication 4. *The average duration of arbitrage opportunities decreases with φ, the likelihood that an arbitrage opportunity is toxic, if $\alpha(4\varphi - 1) \leq 1$ (see the appendix for a proof).*

The direct effect of an increase in φ is to induce the arbitrageur and the market maker to react faster to arbitrage opportunities (holding S_X^* constant, γ^* and λ^* increase in φ; see eq.(15) and (16)). The indirect effect however is that the equilibrium bid-ask spread increases (Implication 1), which reduces traders' incentives to be fast as explained previously. The direct effect dominates when the bid-ask spread is not too large, i.e., when α and φ are not too large so that $\alpha(4\varphi - 1) \leq 1$ (Implication 4). In our data, the number of arbitrage opportunities relative to the total number of trades (a proxy for α) is small and well below 1/3, which is sufficient for this condition to be satisfied for all values of φ. Hence, we expect the duration of arbitrage opportunities to be negatively related to φ.

1.3 Extensions

In this section, we discuss two extensions of the model. For brevity, we omit full derivations of the equilibrium in each case discussed below. They are available upon request. We have checked that our testable implications still hold in each case.

Multiple arbitrageurs. The baseline model features a single arbitrageur. We have also considered the case with $M > 1$ arbitrageurs. This case is more difficult to analyze because then arbitrageurs choose their speed to beat both the market maker and other arbitrageurs. However, the implications of the model still hold in this case. In fact competition among arbitrageurs reinforces dealers' exposure to the risk of trading at stale quotes when a toxic arbitrage occurs because the likelihood (π^*) that the market maker cannot cancel his quotes fast enough increases with the number of arbitrageurs. Thus, the competitive bid-ask spread increases with the number of arbitrageurs (M).

Shocks to market maker X's valuation. In the baseline model, arbitrage opportunities are only due to shifts in market maker Y's valuation (i.e., always originates in asset Y). Suppose instead that a shock to a market maker's valuation at $t = 0$ can affect either market maker X or Y with equal probabilities. Under these assumptions, markets for assets X and Y are perfectly symmetric. Thus, in equilibrium bid-ask spreads are identical in each asset and strictly positive ($S_Y^* = S_X^* > 0$). In this case, there is no closed form solution for the competitive bid-ask spread.[21] The equilibrium can be solved numerically, however. Numerical simulations show that our predictions still hold in this more general case.

2. Data and variables construction

In the rest of the paper, we provide evidence supporting our predictions using data on order submissions and triangular arbitrage opportunities in the foreign exchange (FX) market. In this section, we describe our data and we define the main variables used in our tests. Empirical findings are reported in Section 3.

2.1 Data

We use data from Reuters D-3000 from January 2, 2003 to December 30, 2004 for three currency pairs: US dollar/euro (dollars per euro; hereafter USD/EUR), US dollar/pound sterling (USD/GBP), and pound sterling/euro (EUR/GBP). These pairs account for 60 percent of all FX spot transactions at the time of our sample (see Bank for International Settlements 2005). Our sample has 485 days after excluding weekends and certain holidays (as in, for instance, Andersen et al. 2003) because trading activity is considerably lower during these days.

Reuters D-3000 is an electronic limit order book market in which foreign exchange dealing

[21] In this more general case, π^* is a non-linear decreasing function of the bid-ask spread (in the baseline model, π^* only depends on ρ for any value of the spread). As π^* is non-linear in the bid-ask spread, market makers' expected profits are non linear in their spread and competitive bid-ask spreads solve a cubic polynomial.

banks ("FX dealers") can post quotes or hit quotes posted by other dealers. In 2003-2004, Reuters D-3000 has the dominant market share in the dollar-sterling and euro-sterling pairs while its competitor, EBS (Electronic Broking Services), is dominant in the dollar-euro pair.[22] For our tests, we exclusively focus on triangular arbitrage opportunities within Reuters D-3000. When an arbitrageur exploits a toxic arbitrage opportunity in Reuters D-3000, it inflicts a loss on market makers with stale quotes on this platform. Hence, quotes in Reuters D-3000 should reflect this risk, as predicted by our model.

For each order submitted to Reuters D-3000, our dataset reports (i) the currency pair in which the order is submitted; (ii) the order type (limit or marketable) and its direction (buy or sell); (iii) the size of the order (in millions of the base currency); (iv) its price for a limit order; and (v) the time at which the order is entered with an accuracy of one-hundredth of a second. These data enable us to measure accurately the duration of triangular arbitrage opportunities in our sample, identify their nature (toxic/non-toxic), and observe whether they terminate with an arbitrageurs' trade or a quote update. This is key for our tests (see below).

2.2 Toxic and non-toxic triangular arbitrage opportunities

Let $A_t^{i/j}$ be the amount of currency i required, at time t, to buy one unit of currency j and $B_t^{i/j}$ be the amount of currency i received for the sale of one unit of currency j. These are the ask and bid quotes posted by market makers in currency i versus j at time t. A triangular arbitrage opportunity exists at time t when

$$\hat{A}_t^{i/j} < B_t^{i/j} \qquad \text{or,} \tag{18}$$

$$\hat{B}_t^{i/j} > A_t^{i/j}, \tag{19}$$

where $\hat{A}_t^{i/j} \equiv A_t^{i/k} \times A_t^{k/j}$ and $\hat{B}_t^{i/j} = B_t^{i/k} \times B_t^{k/j}$. In the first case, an arbitrageur can secure a risk free profit, net of bid-ask spread costs, equal to $(B_t^{i/j} - \hat{A}_t^{i/j})$ units of currency i. Indeed, he can first buy $A_t^{k/j}$ units of currency k with currency i for a total cost of $\hat{A}_t^{i/j}$, then use his position in currency k to buy one unit of currency j and finally sell this unit for $B_t^{i/j}$ units of currency i. In the second case, an arbitrageur can secure a risk free profit equal to $(\hat{B}_t^{i/j} - A_t^{i/j})$

[22] See Pierron (2007), Osler (2008), King and Rime (2010), and King et al. (2012) for excellent descriptions of participants, market structure, and recent developments in foreign exchange markets. At the time of our sample, the FX market is a two-tier market. In the first tier, FX dealers trade exclusively with end-users (e.g., hedge funds, mutual funds, pension funds, corporations, etc.). The second-tier is an interdealer market. In this market, dealers can trade (i) bilaterally (by calling each other), (ii) through voice brokers, or (iii) electronic broker systems (e.g., EBS and Reuters D-3000).

units of currency i with a symmetric strategy. We refer to $\hat{A}_t^{i/j}$ and $\hat{B}_t^{i/j}$ as being the "synthetic" quotes for currency i versus j.

Our definition of a triangular arbitrage opportunity accounts for bid-ask spread costs for arbitrageurs. In reality, traders also pay brokerage fees for executing trades on Reuters D-3000. These fees are small and in general well below one basis point (see Chaboud et al. 2014). To account for them, we focus on triangular arbitrage opportunities that deliver a profit of at least 0.2 basis points (bps):

$$\frac{B_t^{i/j} - \hat{A}_t^{i/j}}{\hat{A}^{i/j}} > 0.2 \quad \text{bps}, \quad \text{or} \tag{20}$$

$$\frac{\hat{B}_t^{i/j} - A_t^{i/j}}{\hat{B}_t^{i/j}} > 0.2 \quad \text{bps}. \tag{21}$$

Our empirical findings are qualitatively unchanged if we set a higher bar (1 basis point) for the profitability of triangular arbitrage opportunities.

Our first step is to identify triangular arbitrage opportunities in our data, measure their duration, and record how they terminate. To this end, we proceed as follows (see the appendix for a numerical example).

- Starting from a state in which there is no triangular arbitrage opportunity (i.e., a state in which Conditions (20) and (21) do not hold), we record the latest best bid and ask prices for the three currency pairs each time a new limit order is submitted and we check whether a triangular arbitrage opportunity exists (using Conditions (20) and (21)). If this is the case we deduce that the limit order arrival created the opportunity and we record the order arrival time, t_0, as the time at which the arbitrage opportunity begins.

- We then record the time t_1 at which the arbitrage opportunity disappears and define the duration of the arbitrage opportunity as $(t_1 - t_0)$. We also record whether the arbitrage opportunity terminates with a trade from an arbitrageur or quote updates by market makers (the only two ways in which an arbitrage opportunity can terminate).

By definition, a triangular arbitrage opportunity happens in all three currency pairs at the same time. However, the opportunity is triggered by a new limit order submitted in one specific pair, which we call the "initiating pair."[23] The number of arbitrage opportunities in our data

[23] The arrival of a market order cannot create an arbitrage opportunity. For instance, suppose that $A_t^{i/j} > \hat{B}_t^{i/j}$ and $\hat{A}_t^{i/j} > B_t^{i/j}$, so that, at time t, there is no arbitrage opportunity. If a buy (resp. sell) market order arrives

is therefore the number of times a new limit order creates a triangular arbitrage opportunity. Using this methodology, we count $172,044$ triangular arbitrage opportunities in our sample.

In a second step, we sort all triangular arbitrage opportunities into two groups labeled toxic (i.e., due to an asynchronous reaction of rates to information arrival) and non-toxic (i.e., due to a transient price pressure in one rate). To this end, we proceed as in Schultz and Shive (2010). That is, for each triangular arbitrage opportunity in our sample, we compare the exchange rate for the initiating currency pair when the arbitrage opportunity begins (time t_0) and when it terminates (time t_1). If we observe a reversal of the rate toward its level just before the opportunity starts then we consider that the arbitrage opportunity is due to a transient price pressure. Hence, we classify it as non-toxic. Otherwise, we consider that the arbitrage opportunity is due to asynchronous price adjustments in the rates of the three currency pairs and we classify it as toxic.

[Insert Figure 1 about here.]

Figure 1 illustrates this methodology for four arbitrage opportunities. The solid and dashed lines in Panels A and B show the evolution of actual and synthetic bid and ask quotes during these opportunities. In Panel A, actual and synthetic quotes of the currency pair initiating the arbitrage (EUR/GBP) opportunity shift to a new level when the arbitrage opportunity terminates. The pattern is consistent with the arrival of fundamental information (e.g., macroeconomic news or headlines news on Reuters). We classify these opportunities as toxic. In contrast, in Panel B, only the quotes of the initiating pair (EUR/USD) change during the arbitrage opportunity. Moreover, when the arbitrage opportunity terminates, these quotes revert to their initial level. This pattern (reversal and the absence of changes in the synthetic quotes) is consistent with a transient price pressure in the initiating pair. Accordingly, we classify these arbitrage opportunities as non-toxic.

[Insert Figure 2 about here]

Using this methodology, we identify $83,488$ toxic arbitrage opportunities, i.e., 48% of all opportunities. Panel A of Figure 2 shows the time-series of the daily number of (a) all triangular

in the market for currency i versus j, it executes at $A_t^{i/j}$ (resp., $B_t^{i/j}$) and possibly higher (resp., lower) prices if the size of the order is larger than the quantity available at $A_t^{i/j}$ (resp. $B_t^{i/j}$). Thus, just after the arrival of a buy (sell) market order (at time t^+), the new best offer (bid) price is $A_{t^+}^{i/j} \geq A_t^{i/j} > \hat{B}_t^{i/j}$ ($B_{t^+}^{i/j} \leq B_t^{i/j} < \hat{A}_t^{i/j}$), so that no arbitrage opportunity exists if there is none at date t. In contrast, if a market maker posts a new ask price $A_{t^+}^{i/j}$ at time t^+ such that $A_{t^+}^{i/j} < \hat{B}_t^{i/j}$ then he creates an arbitrage opportunity. The same is true for orders arriving in other currency pairs and affecting the synthetic quotes.

arbitrage opportunities (light grey line) and (b) toxic arbitrage opportunities (black line) in our sample. There is substantial variation in the number of arbitrage opportunities per day and the number of these opportunities that are toxic. On average, per day, there are 112 toxic triangular arbitrage opportunities (standard deviation (s.d.)=49.62) and 108 non-toxic arbitrage opportunities (s.d.=39.22).

Panel B of Figure 2 shows average intra-day patterns in the number of arbitrage opportunities. The bulk of the activity for currency pairs in our sample occurs from 7:00 GMT when European markets open until 17:00 GMT when European markets close. Not surprisingly, most arbitrage opportunities occur during this period, with peaks when trading activity in the U.S. and in Europe overlap (13:00 to 17:00). Hence, we only retain observations from 7:00 to 17:00 GMT for the variables used in our tests (see the next section).

2.3 Variables of interest and summary statistics

Our tests focus on the effects of (i) the likelihood that an arbitrage opportunity is toxic (φ in the model) and (ii) the likelihood that a toxic arbitrage opportunity terminates by an arbitrageur's trade (π^* in the model) on illiquidity and the duration of arbitrage opportunities. We explain below how we measure empirically these variables.

We use four different measures of illiquidity. Our first three measures are daily averages of the (i) the quoted bid-ask spread, $spread_{it}$ (for currency pair i on day t), (ii) the effective spread ($espread_{it}$), and (iii) a measure of the price impact of trades ($adv\_selection_{it}$). The effective bid-ask spread for a particular transaction is defined as twice the buy-sell indicator for the transaction (+1 for a buy and −1 for a sale) multiplied by the difference between the transaction price and the prevailing mid-quote. The price impact of a trade is defined as the change in the mid-quote following the trade over one minute following the transaction times the buy-sell indicator. The average price impact over day t in currency i is a measure of losses due to adverse selection for liquidity suppliers in this currency on this day (see, for instance, Hendershott et al. 2011). According to our model, an increase in dealers' exposure to toxic arbitrage should raise illiquidity because it increases adverse selection costs for dealers.

In our model, dealers react to an increase in their exposure to toxic arbitrage by increasing their bid-ask spread. In reality, they might also reduce the size of their limit orders (quoted depth). Our fourth measure of illiquidity, $slope_{it}$, accounts for this possibility. It is defined as the equally weighted average of two ratios: (i) the difference between the second and first best ask prices divided by the number of shares offered at the best ask price and (ii) the difference between the first and second best bid prices divided by the number of shares offered at the best bid price for currency i on day t. Hence, $slope_{it}$ is higher when the number of shares offered at the best quotes is smaller and the second best prices in the book are further away from the best quotes. A higher $slope_{it}$ implies that the limit order book is thinner for currency i on day t.

The duration of arbitrage opportunities on day t ($E(D)$ in the model) is measured by $duration_t$, which is the average duration of all triangular arbitrage opportunities on day t in our sample.

We now turn to our two main explanatory variables: the arbitrage mix, φ, and arbitrageurs' relative speed, π. On each day t, we measure the likelihood that an arbitrage opportunity is toxic, φ_t, by the number of toxic arbitrage opportunities divided by the total number of arbitrage opportunities on day t. That is,

$$\varphi_t = \frac{\text{No. of toxic arbitrage opportunities on day } t}{\text{No. of all arbitrage opportunities on day } t}. \tag{22}$$

Similarly, on each day t, we measure the likelihood, π_t^{tox}, that an arbitrageur is fast enough to exploit a toxic arbitrage opportunity by the number of toxic arbitrage opportunities that terminate with a trade divided by the total number of arbitrage opportunities on day t:

$$\pi_t^{tox} = \frac{\text{No. of toxic arbitrage opport. that terminate with a trade on day } t}{\text{No. of toxic arbitrage opportunities on day } t}. \tag{23}$$

The likelihood of occurrence of an arbitrage opportunity, α, and the size of arbitrage opportunities, σ, also affect illiquidity according to the model (see eq.(13)). We therefore control for these variables in our tests. We measure α on day t by the number of all arbitrage opportunities on day t divided by the total number of trades on this day (denoted α_t). For σ, we use the average absolute percentage difference between the mid-points of the actual quotes on the one hand and synthetic quotes on the other hand for all currency pairs in our sample at the time of each toxic arbitrage opportunity on day t.[24] We denote this difference on day t by σ_t^{tox}.

[Insert Table 2 here]

Table 2 presents summary statistics for the main variables in our analysis. Panel A reports

[24]Formally, suppose that a toxic arbitrage opportunity occurs at time τ on day t. Let $f_{\tau,t}^{i/j} = \frac{A_{\tau,t}^{i/j}+B_{\tau,t}^{i/j}}{2}$ and $\hat{f}_{\tau,t}^{i/j} = \frac{\hat{A}_{\tau,t}^{i/j}+\hat{B}_{\tau,t}^{i/j}}{2}$ be the mid-points of actual quotes and synthetic quotes, respectively, in the market of currency i versus j at this time. We define σ_t^{tox} as the average value of $|(f_{\tau,t}^{i/j} - \hat{f}_{\tau,t}^{i/j})/f_{\tau,t}^{i/j}|$ over all toxic arbitrage opportunities on day t.

that $\alpha_t = 6.5\%$ (s.d= 1.6%) on average, which means that, in a given day, there is one arbitrage opportunity every twenty trades. About half of these opportunities are toxic ($\varphi_t = 50.1\%$ on average). Panels B and C present the characteristics of toxic and non-toxic arbitrage opportunities. Both types of opportunities vanish very quickly: they last on average for about 1.71 seconds (s.d= 0.524) and 1.448 seconds (s.d.= 0.505), respectively. The likelihood that a toxic arbitrage opportunity terminates with a trade (π^{tox}) is 63.1% on average (s.d.= 0.059). For a non-toxic arbitrage opportunity, this likelihood (denoted π^{nontox}) is smaller and equal to 47%.

The average size of a toxic arbitrage opportunity, σ_t^{tox}, is 2.669 basis points (s.d. = 0.529) and, after accounting for trading costs (as in eq. (20) and (21)), the average daily arbitrage profit on a toxic arbitrage opportunity is 0.651 basis points (s.d.= 0.148) per dollar traded. The minimum quoted depth on Reuters is one million of base currency. Thus, the average profit on a toxic triangular arbitrage opportunity is at least \$65.1 per opportunity, i.e., \$7,291.2 per day (since there are 112 opportunities per day on average). Very similar figures are obtained for non-toxic arbitrage opportunities. As a point of comparison, Brogaard et al. (2014) report that, after accounting for trading fees, high frequency traders in their sample earn, in aggregate, \$4,209.15 per stock-day on their market orders in large-cap stocks and much less in small-cap stocks (see Table 4 in Brogaard et al. 2014).

To save space, we report summary statistics for our various measures of illiquidity in Table IA.1.1 (Panels A and B) of the Internet Appendix. Average quoted and effective bid-ask spreads are very tight (between 1 and 3 basis points on Reuters). We also compare the average quoted depth and quoted bid-ask spread in toxic and non-toxic arbitrage opportunities just before these opportunities (see Panel C of Table IA.1.1). Quoted depth is not statistically different between toxic and non-toxic arbitrage opportunities and it varies between 2.5 million of dollar (for USD/EUR) and 3.2 millions dollar (for EUR/GBP). Quoted spreads tend to be slightly higher just before the occurrence of toxic arbitrage opportunities.

Table IA.1.2 in the Internet Appendix reports the correlation of the main variables used in our tests. Consistent with Implication 1, all measures of illiquidity are positively and significantly correlated with φ_t, the fraction of arbitrage opportunities that are toxic. Moreover, as expected, they are also positively and significantly correlated with the size of toxic arbitrage opportunities (σ_t^{tox}) and the frequency (per trade) of these opportunities (α_t). In contrast, the correlation between π^{tox} (our proxy for π^*) and measures of illiquidity is not significantly different from zero. This is not surprising because arbitrageurs should react faster to arbitrage opportunities when trading costs are smaller. This effect works to make π^{tox} higher when bid-ask spreads are smaller, even if π^{tox}'s effect on illiquidity is positive. This highlights the importance of accounting for endogeneity in analyzing the effect of arbitrageurs' relative speed on illiquidity.

The duration of arbitrage opportunities and the various measures of market illiquidity are positively correlated. That is, toxic arbitrage opportunities last longer on average when the market for the three currency pairs is more illiquid. Again this is not surprising because, other things equal, a higher bid-ask spread should induce arbitrageurs to react more slowly to arbitrage opportunities, which eventually results in more persistent arbitrage opportunities. Our Implication 3 however shows that a positive shock to arbitrageurs' relative speed can simultaneously increase bid-ask spreads while reducing the duration of arbitrage opportunities.

3. Empirical evidence

3.1 Is liquidity sensitive to the arbitrage mix and arbitrageurs' relative speed?

We first test whether illiquidity increases in (i) the likelihood that an arbitrage opportunity is toxic, φ_t, and (ii) the likelihood that an arbitrage opportunity terminates with an arbitrageur's trade, π_t^{tox}, as predicted by Implications 1 and 2. For this, we regress each measure of illiquidity for each currency pair, $illiq_{it}$, on π_t^{tox}, φ_t, and various controls (stacked in vector X_{it}), for time-varying market conditions. That is, we estimate:

$$illiq_{it} = \omega_i + \xi_{t,m} + b_1\pi_t^{tox} + b_2\varphi_t + b^{'}X_{it} + \varepsilon_{it}, \tag{24}$$

where ω_i is a currency pair fixed effect and $\xi_{t,m}$ is a monthly fixed effect (a dummy equal to one if day t is in month m). Coefficients b_1 and b_2 should be positive according to Implications 1 and 2. All our variables are measured at the daily frequency. There are two days without any arbitrage opportunity, which prevents us from computing π^{tox} on these days. Hence we exclude them from our sample and we eventually conduct our tests with a sample of 483 days (i.e., 1449 currency-day observations).

The vector of control variables, X_{it}, includes α_t and σ_t^{tox} (defined in Section 2.3) and additional control variables. First, we use currency-specific controls known to be correlated with illiquidity, namely (i) $trsize_{it}$, the average trade size in currency i on day t, (ii) vol_{it}, the realized volatility (the sum of squared five minutes mid-quote returns) in currency i on day t, and (iii) $nrorders_{it}$, the number of orders (new limit and market orders as well as limit order updates) in currency i on day t. As in Pasquariello (2016), we also control for the TED spread (denoted ted_t on day t), i.e., the difference between the LIBOR and the T-Bill rate because variations in funding costs can affect liquidity and the duration of arbitrage opportunities (seeBrunnermeier et al. 2008). Market-wide sources of variations in liquidity should affect in the same way the liquidity of a currency pair on Reuters and its competitor, EBS. Thus, in estimating eq.(24) for a particular illiquidity measure on Reuters (say, the quoted bid-ask spread), we include its EBS counterpart (denoted $illiq_{it}^{EBS}$ for currency i on day t) in our set of controls.[25] Finally, we use the number of days since the beginning of the sample to control for a possible trend in our illiquidity measures.

In the model, the likelihood that a toxic arbitrage terminates with an arbitrageur's trade (π_t^{tox}) is endogenous and simultaneously determined with the bid-ask spread. Thus, to identify the effect of π_t^{tox}, we use an instrumental variable (IV). In July 2003, Reuters D-3000 introduced a new functionality, "Reuters AutoQuote API" (Application Programming Interface), allowing traders to automate order submission instead of manually typing trading instructions. This functionality marked the beginning of algorithmic trading on Reuters by enabling traders to input Reuters datafeed in their algorithms and let these trade accordingly. The order-to-trade ratio (the number of orders to the number of trades) is often used as an indicator of algorithmic trading activity (see Hendershott et al. 2011). Figure 3 shows that this ratio for Reuters D-3000 experiences a significant increase in July 2003, suggesting that some traders quickly took advantage of Autoquote API to automate their trading decisions.

[Insert Figure 3 about here.]

Two conditions must be satisfied for AutoQuote API (henceforth "AutoQuote") to serve as a valid instrument for π_t^{tox}, the likelihood that a toxic arbitrage terminates by an arbitrageur's trade. First, AutoQuote should have a clear effect on this likelihood. Second, it should satisfy

[25] Our EBS data are identical to those for Reuters D-3000, except that all orders and trades on EBS occurring within the same second receive the same time stamp. Thus, EBS data cannot be used to accurately measure when a triangular arbitrage opportunity (across trading systems or within EBS) starts, and when and how it terminates. For example, suppose two market orders and two limit orders are submitted in a second in which an arbitrage opportunity occurs and that the arbitrage starts and terminates within this second. EBS data do not allow us to identify whether the arbitrage terminates due to a marketable order (a trade) or a limit order (a quote update). Hence, we cannot compute π^{tox} and φ using EBS data.

the "exclusion restriction," i.e., the introduction of Autoquote should not be correlated with the error term in eq.(24).

The first condition is very likely to be satisfied because AutoQuote reduces the time cost of reacting quickly to arbitrage opportunities for market participants. Thus, it should affect arbitrageurs relative cost of speed (ρ in the model) and therefore arbitrageurs relative speed (π in the model) since the latter increases with the former in equilibrium (see eq.(14)). For identification, it is not important whether the sign of this effect is positive or negative: it suffices that the effect exists. However, anecdotal evidence suggests that Autoquote was, at least initially, predominantly use by arbitrageurs. For instance, Pierron (2007) writes that Autoquote API: "*allows a full benefit from algorithmic trading, since it enables the black box to route the order to the market with the best prices and potential arbitrage across markets despite the fragmentation of the various pools of liquidity in the FX market.*" Similarly, Chaboud et al. (2014) note that (on p.2058): "*From conversations with market participants, there is widespread anecdotal evidence that in the very first years of algorithmic trading in this [FX] market, a fairly limited number of strategies were implemented with triangular arbitrage among the most prominent.*" Thus, we expect Autoquote to increase the likelihood that a toxic arbitrage terminates by an arbitrageur's trade. The first stage of the IV regression (see below) confirms this conjecture.

The second condition–the exclusion restriction–requires that Autoquote should affect illiquidity only through its effect on π^{tox}, after controlling for other variables appearing in eq.(24). One concern is that the introduction of Autoquote might coincide with other contemporaneous shocks to the liquidity of the currencies in our sample. However, to bias our estimate for the effect of π^{tox}, these shocks should be unrelated to our control variables (e.g., funding liquidity shocks and volatility). Moreover, they should be specific to the Reuters trading platform since we control for systematic variations in the liquidity of the currency pairs in our sample by including measures of illiquidity for the same pairs on EBS ($illiq_{it}^{EBS}$).

Another threat to identification is that Autoquote directly influences the costs of market making or rents earned by market makers. In particular, automation of order entry might reduce dealers' order processing costs or increase competition among dealers by triggering entry of new participants. Our control for the number of order submissions on Reuters ($nrorders_{it}$) should capture, to some extent, the effect of new entry since the volume of order submissions should increase with the number of participants. In any case, if they are present, these effects imply that Autoquote should be associated with a decrease in bid-ask spreads. Hence, they should bias downward our estimate of the effect of π^{tox} on illiquidity and make it more difficult for us to detect a positive effect. Moreover, our measure of adverse selection costs for dealers ($adv\_selection_{it}$) is immune to these problems because this component of bid-ask spreads is not affected by order processing costs and market power (see, for instance, Stoll 2000). The possibility remains that Autoquote affects adverse selection costs through channels other than the risk of trading at stale quotes when a toxic arbitrage opportunity arises (the source of adverse selection in our model). We address this concern in Section 3.3.

Insert Table 3 About Here

Table 3 reports estimates of eq.(24) using the introduction of Autoquote API as an instrument. The first stage of the IV regression is:

$$\pi_t^{tox} = \omega_i + \xi_{t,m} + a_1 AD_t + a_2\varphi_t + a' X_{it} + u_t, \tag{25}$$

where AD_t is our instrument (a dummy equal to 1 after July 2003 and zero before) and X_{it} is the same set of control variables as in eq.(24).[26] In our estimation, we account for time series autocorrelation and heteroscedasticity in residuals in computing standard errors.

Estimates of the coefficients of the first stage (eq.(25)) are reported in Columns (1), (3), (5) and (7) of Table 3 (one column per illiquidity measure). They are very similar across all illiquidity measures.[27] The coefficient (a_1) on the dummy variable, AD_t, is statistically significant and equal to about 0.035. The instrument is not weak since the F statistics is around 28 in all specifications.[28] Overall, as we expected, the first stage regression indicates that AutoQuote had a positive effect on arbitrageurs' relative speed. This is consistent with Chaboud et al. (2014), who note (p.2067) that triangular arbitrage opportunities in their sample disappear mainly through "*computer hitting existing quotes, with most of these quotes posted by human traders.*"

[26] We have not been able to retrieve the exact day of the introduction of Autoquote API in July 2003. Thus, AD_t takes the value one only starting in August 2003. This is unlikely to affect our results since the adoption of Autoquote API certainly took a few weeks.

[27] The first stage is specific to each illiquidity measure because our set of control variables include $illiq_{it}^{EBS}$, which is specific to the illiquidity measure used as dependent variable.

[28] One should perform the Sargan-Hansen test of overidentifying restrictions routinely in any overidentified model estimated with instrumental variables to check whether the excluded instruments are appropriately independent of the error process. This is done if and only if an equation is overidentified. We do not carry out the test because we have a just-identified case in our setup.

We now turn to our main variables of interest, i.e., the estimates for coefficients b_1 and b_2 in eq.(24). For all illiquidity measures, we find that illiquidity is higher when the fraction of arbitrage opportunities that are toxic is higher, as predicted by Implication 1. This effect is both statistically and economically significant. For instance, for the quoted bid-ask spread, we find that $b_2 = 0.875$ (t-stat=3.53) and for the adverse selection cost, we find that $b_2 = 0.428$ (t-stat=4.11). Thus, an increase in φ_t by a one standard deviation (i.e., 0.06) is associated with an increase of 0.0257 basis points in the adverse selection cost for the currencies in our sample (i.e., about 40% of the average quoted spread in our sample).

As predicted, the effect of π^{tox} on illiquidity, b_1, is also positive and statistically significant for all measures of illiquidity. For instance, for the quoted bid-ask spread, we find that $b_1 = 6.341$ (t-stat=4.55) and for the adverse selection cost, we find that $b_1 = 2.223$ (t-stat=3.43). Thus, a 1% increase in π_t^{tox} raises the quoted bid-ask spread ($spread_{it}$) by 0.0634 basis points (t-stat = 4.55). The economic size of this effect is significant as well since it represents about 3% of the average bid-ask spread (about 2 basis points) for currencies in our sample.

Another way to evaluate the economic significance of this finding is to consider the effect of an increase of 1% in π_t^{tox} on daily trading costs. The average trade sizes (in dollar) for the currencies considered in our sample are \$2.404 million in GBP/USD, \$1.667 million in EUR/USD, and \$1.839 million in EUR/GBP. Moreover, the average number of transactions per day is 4.751 in GBP/USD, 2.399 in EUR/USD, and 2.876 in EUR/GBP. Hence, according to our estimates, a 1% increase in π_t^{tox} raises total daily trading costs by $0.06341 bps \times (\$2.404 \times 4.751 + \$1.667 \times 2.399 + \$1.839 \times 2.876) = \$131,319$ for the three markets in total. Thus, even a small increase in arbitrageurs' speed can be rather costly for other market participants given the large volume of trade in currency markets.

Other control variables in eq.(24) are signed as expected. For instance, daily changes in illiquidity are positively and significantly related to the size of toxic arbitrage opportunities (σ_t^{tox}) and the likelihood of occurrence of an arbitrage opportunity (α_t). They are also positively associated with realized volatility and funding constraints.

In the Internet Appendix (see Table IA.2.1 in this appendix), we report estimates of eq.(24) with Ordinary Least Squares (OLS) instead of the instrumental variable (IV) approach. The OLS estimate for the effect of π^{tox} (i.e., b_1) is positive. However, it is in general smaller than the IV estimate and its statistical significance is weaker. This is consistent with the fact that traders' relative speed is endogenous and jointly determined with illiquidity.

Hendershott et al. (2011) find that algorithmic trading improves liquidity, mainly because it reduces the adverse selection cost component of the bid-ask spread (as measured by price impact). At first glance, this finding seems to contradict ours. However, our tests do not focus on the effect of the same variables on liquidity and our results are therefore not directly comparable. Specifically, we study the effects of the arbitrage mix (φ) and arbitrageurs' relative speed (π^{tox}) while Hendershott et al. (2011) study the effect of algorithmic trading measured by AT, defined as the number of electronic messages normalized by trading volume for NYSE stocks. Importantly, Hendershott et al. (2011) emphasize that variations in AT are mainly driven by variations in limit order submissions and cancellations. Thus, AT mainly picks up algorithmic liquidity provision, not algorithmic arbitrage since algorithmic arbitrageurs use market orders (take liquidity), not limit orders (see Chaboud et al. 2014). For this reason, the AT variable in Hendershott et al. (2011) is likely to be positively associated with dealers' relative speed and therefore inversely related to π^{tox}. Thus, according to our model, an increase in AT should reduce the adverse selection cost component of the bid-ask spread, which is indeed Hendershott et al. (2011))'s finding. This discussion suggests that the correlation between proxies for algorithmic trading and liquidity providers' relative speed is key to interpret how these proxies are associated with liquidity measures.[29]

3.2 Duration of arbitrage opportunities

As explained in the previous section, the introduction of Autoquote reduces arbitrageurs' cost of speed. Thus, according to Implication 3, Autoquote should be associated with shorter arbitrage opportunities. Moreover, Implication 4 predicts that the duration of arbitrage opportunities should be shorter on days in which the fraction of toxic arbitrage opportunities is higher. To test these implications, we estimate the following equation:

$$\log(duration_t) = \omega_i + \xi_{t,m} + c_1 AD_t + c_2 \varphi_t + c^{'} X_{it} + u_t, \tag{26}$$

where $duration_t$ is the average duration of arbitrage opportunities on day t and X_{it} is the same vector of control variables as in eq.(24) (where $illiq^{EBS}$ is measured by the EBS quoted spread).

[29] This point might explain why various papers reach different conclusions about the effect of trading speed on liquidity. For instance, Hendershott and Moulton (2011) find empirically that the increase in speed of execution for marketable orders (a decrease in liquidity providers' relative speed) results in an increase in illiquidity while Brogaard et al. (2015) find that an increase in liquidity providers' speed of access to the market (through upgrades in co-location services) improves liquidity.

We do not control for π_t^{tox} in this case because, in theory, the duration of arbitrage opportunities is affected by shocks (i.e., AD_t in our test) to arbitrageurs' cost of speed directly, not through π_t^{tox} (see eq.(17)). Implications 3 and 4 predict that c_1 and c_2 should be strictly positive.

[**Insert Table 4 about here**]

Estimates of eq.(26) are reported in Table 4. In Column (1), the dependent variable is the log of the duration of toxic arbitrage opportunities, while in Column (2) the dependent variable is the log of the duration of all arbitrage opportunities. Consistent with Implication 3, Autoquote (AD) is associated with shorter arbitrage opportunities. For instance, in Column (1), $a_1 = -0.067$ (t-stat$=-3.68$). This point estimate implies that the introduction of Autoquote reduced the duration of toxic arbitrage opportunities by about 6.7%.[30] Relative to the average duration of toxic arbitrage opportunities (1710 milliseconds; see Table 2), this implies a drop of about 115 milliseconds (6.7% $\times$ 1.710) in the duration of toxic arbitrage opportunities after the introduction of Autoquote. Very similar estimates are obtained when we use the average duration of all arbitrage opportunities.

Moreover, as predicted by Implication 4, we find that, on average, the duration of arbitrage opportunities is shorter when the fraction of arbitrage opportunities that are toxic, φ, is higher. For instance, in Column 1, $a_2 = -0.934$ (t-stat$=-7.91$). Thus, a one standard deviation increase in the fraction of arbitrage opportunities that are toxic (i.e., an increase of 0.06, see Table 2) is associated with a 5.6% drop in the duration of toxic arbitrage opportunities.

3.3 Exposure to toxic arbitrage or other forms of adverse selection?

As explained previously, one might be concerned that Autoquote affects dealers' adverse selection costs through other channels than dealers' exposure to toxic arbitrage. Another concern is that our measures of arbitrage toxicity (φ and π^{tox}) might just pick the effect of more traditional measures of adverse selection. To address these concerns, we add four different measures of adverse selection in the set of control variables used in eq.(24) and study whether this alters our conclusions regarding the effects of φ and π^{tox} on illiquidity.

Our first measure is a dummy variable equal to one on days with "influential" macroeconomic announcements for currency pairs in our sample (i.e., about the EMU, U.K., and

[30]The coefficient on a particular control variable in eq.(26) measures the percentage change in the duration of arbitrage opportunities following a one unit change in this control variable because our dependent variable is the log of duration.

U.S.).[31] Indeed, Green (2004) finds empirically that the release of macro-economic news is associated with an increase in informed trading. Overall, there are 242 days with at least one macro-announcement in our sample.

In addition, Pasquariello and Vega (2007) shows that adverse selection costs are higher when informed traders' beliefs are more heterogeneous. Thus, as a second proxy for adverse selection, we use a measure of heterogeneity in informed traders' beliefs, which, as in Pasquariello and Vega (2007), is based on the dispersion of professional forecasts about macro-economic announcements. Specifically, Bloomberg asks a panel of experts to provide forecasts about the next macro-economic announcement one week before its release. Using these forecasts, for, say, the n^{th} announcement of type j, we compute the standard deviation, SD_{jn}, of experts' forecasts (across experts) and compute

$$std_{jn}^{macro} = \frac{SD_{jn} - \mu(SD_{jn})}{\sigma(SD_{jn})}, \tag{27}$$

where $\mu(SD_{jn})$ and $\sigma(SD_{jn})$ are, respectively, the mean and standard deviation of SD_{jn} over all announcements of type j. Then, as in Pasquariello and Vega (2007), we assume that this measure of heterogeneity in professional forecasts about the announcement of type j is constant every day in-between announcements of this type and we measure the heterogeneity in informed traders' beliefs on day t by std_t^{macro}, i.e., the average of std_{jn}^{macro} on day t across all announcements.

Third, following Easley et al. (2008) and Holden and Jacobsen (2014), we also consider the average absolute daily order imbalance across the three currency pairs in our sample as another measure of adverse selection. Specifically, for each currency i on day t, we compute

$$|oib_t^i| = \frac{|buys_t^i - sells_t^i|}{buys_t^i + sells_t^i}, \tag{28}$$

where $buys^i$ ($sells^i$) is the number of buy (sell) market orders for currency pair i on day t. We then define oib_t as the average of $|oib_t^i|$ across the three currency pairs on day t and use it as another measure of adverse selection. Importantly, in computing $|oib_t^i|$, we exclude buy and sell market orders that terminate an arbitrage opportunity to mitigate the concern that $|oib_t^i|$ might

[31] Influential macro-economic announcements are defined as in Pasquariello and Vega (2007) or Rime et al. (2010). These are non-farm payroll (US), leading indicators (US), retail sales (US), new home sales (US), consumer confidence index (US), unemployment claims (US), NAPM index (US), current account (EMU), industrial production (EMU), CPI (EMU), retail sales (EMU), trade balance (UK), and GDP growth (UK).

reflect arbitrageurs' activity.

Last Easley et al. (2012) advocates the use of VPIN ("volume-synchronized probability of informed trading") as a measure of high-frequency adverse selection. Following Easley et al. (2012), in each trading day t, we group successive trades into 50 equal volume buckets of size v_t^i, where v_t^i is equal to the trading volume in day t for currency i divided by 50.[32] The $vpin_t^i$ metric for currency pair i and day t is then

$$vpin_t^i = \frac{\sum_{\tau=1}^{50} \left| v_\tau^{i,s} - v_\tau^{i,b} \right|}{50 \times v_t^i}, \tag{29}$$

where $v_\tau^{i,b}$ and $v_\tau^{i,s}$ are the amount of base currency purchased and sold, respectively, within the τ^{th} bucket for currency pair i. As for the order imbalance measure and for the same reason, we exclude arbitrageurs' trades when computing $vpin_t^i$.

[Insert Table 5 about here]

Table 5 reports the result of the IV regression when we control for days with macro-economic announcements (D^{macro}), dispersion of professional forecasts (std^{macro}), absolute order imbalance ($|oib|$) and VPIN ($vpin$). In general, these variables are positively related to our various measures of illiquidity but this relationship is not statistically significant. More importantly for our purpose, the effects of φ_t and π_t^{tox} on illiquidity remain positive and significant for all measures of illiquidity. Furthermore, estimates of the effect of these variables are very similar to those reported in Table 3. Hence, φ_t and π_t^{tox} contain information about market makers' exposure to adverse selection, that is not subsumed in other measures of adverse selection.

3.4 Robustness

We conduct additional robustness checks, which, for brevity, are only reported in the Internet Appendix. First, we estimate eq.(24) and eq.(26) at the hourly frequency rather than at the daily frequency. Results are qualitatively similar and statistically stronger due the increase in the number of observations (see Table IA.3.1 in the Internet Appendix). We also estimate eq.(24) and eq.(26) with other proxies for funding liquidity shocks than the TED spread, namely (i) aggregate primary dealer repo positions as in Adrian and Shin (2010), (ii) difference between the

[32]For robustness, we have also grouped trades into different volume buckets of 30, 75, and 100 in the construction of VPIN. Results reported below are not sensitive to the size of grouping.

three-month commercial paper rate and the three-month Treasury Bill rate as in Krishnamurthy (2002) and (iii) a measure of noise in U.S. Treasury yields as constructed by Hu et al. (2013). Our findings are not sensitive to how we measure funding liquidity shocks (see Table IA.3.2). In Table IA.3.3, we use the quoted depth (in million of dollar of the base currency) as another measure of liquidity. We again find that increases in π_t^{tox} and φ_t are associated with a drop in liquidity (i.e., a reduction in quoted depth). The effect is statistically significant only for π_t^{tox}. However, as an increase in φ_t is associated with a larger quoted bid-ask spread, liquidity unambiguously drops when φ_t is higher. Finally, in Table IA.3.4, we measure adverse selection costs with 5-minutes price impacts. Results with this measure are identical to those obtained with 1-minute price impacts as a measure of adverse selection costs.

By construction, the arbitrage mix, φ_t, is positively related to the number of toxic arbitrage opportunities and negatively related to the number of non-toxic arbitrage opportunities. To analyze separately the role of each variable, we replace φ_t by the logarithm of the number of toxic arbitrage opportunities ($log(narb^{tox})$) and the logarithm of the number of non-toxic arbitrage opportunities ($log(narb^{nontox})$) in our regressions. We find (see Table IA.3.5) that the number of toxic arbitrage opportunities has a significant and positive effect on illiquidity, consistent with the notion that toxic arbitrage opportunities increase illiquidity. In contrast, the number of non-toxic arbitrage opportunities has no significant effect on illiquidity. Moreover, the duration of arbitrage opportunities is negatively related to the number of toxic arbitrage opportunities, consistent with the fact that both arbitrageurs and market makers should monitor the market more closely when there are more toxic arbitrage opportunities.

Liquidity providers might be more sensitive to surprises in the fraction of toxic arbitrage opportunities in a given day than to the expected level for this fraction. For instance, a sudden unexpected increase in the number of toxic arbitrage opportunities might lead dealers to significantly cut on liquidity provision. To test for this possibility, we decompose φ_t into an anticipated component ($\varphi_t^{anticipated}$) and an unanticipated component ($\varphi_t^{surprise}$). For this decomposition, we estimate an autoregressive model for φ_t with 20 lags (to capture persistence in the level of φ_t) and additional lagged control variables for the three currencies in our sample. On day t, $\varphi_t^{anticipated}$ is the predicted value of φ_t according to the forecasting model and $\varphi_t^{surprise}$ is the residual of this model. We then re-estimate eq.(24) and eq.(26) allowing the coefficients on the anticipated and the unanticipated components of φ_t to be different. We report estimates obtained with this approach (including those for the forecasting model for φ_t)

in Table IA.3.6 in the Internet Appendix. We find that both the anticipated and unanticipated components of φ_t are positively associated with all measures of illiquidity. Coefficients on both components are statistically significant at least at the 10% level, with, in general, a stronger statistical and economic significance for the anticipated component. Moreover, the duration of arbitrage opportunities is significantly smaller when either components of φ_t are higher.

Our classification of arbitrage opportunities into toxic and non-toxic arbitrage opportunities assumes that price reversals associated with non-toxic arbitrage occur immediately after the arbitrage terminates. If instead price reversals take more time, our approach might mistakenly classify an arbitrage opportunity as toxic while in fact it is not. Hence, as a robustness check, we lengthen the window of observation following an arbitrage opportunity to classify it as toxic or non-toxic. Specifically, we classify an arbitrage opportunity as toxic only if the three quote updates after its termination do not lead to a price reversal. As quote updates are very frequent in our data (the average inter-quote duration is about 1.18 second), this time window should be long enough for rates to revert after the termination of an arbitrage opportunity due to a price pressure, while alleviating the concern that the reversal is due to other confounding factors. With this approach, the number of toxic arbitrage opportunities drops to $68,053$ (i.e., 40.13% of all arbitrage opportunities). However, the results of our tests with this new classification are qualitatively similar to those obtained previously (see Table IA.3.7 in the Internet Appendix).

3.5 External validity

For our tests, we are constrained by data availability. Our data from Reuters D-3000 allow us (i) to accurately measure when and how an arbitrage opportunity terminates and (ii) to instrument arbitrageurs' relative speed with the introduction of Autoquote API. We are not aware of other datasets with these features, which are necessary for our tests.[33] One concern is that trading was slower during our sample period than it is today. This raises the question of whether our results are relevant for today's markets. We think that this is the case for three reasons.

First, the arbitrage opportunities in our sample are very short lived (see Table 2). Thus, at the time of our sample, speed was already of paramount importance for exploiting these

[33] In particular, arbitrage typically take place between assets traded in different platforms (e.g., a stock that trades on multiple trading platforms). This creates problems for studying arbitrage opportunities at the high frequency because the clocks on which different platforms report orders and trades are often not perfectly synchronised. This problem does not arise in our data because we focus on triangular arbitrage opportunities within the same trading platform (Reuters D-3000).

opportunities as it is in today's markets. Second, Reuters D-3000 is an electronic limit order market. Thus, its market structure is very similar to that in which today's fast traders operate. Last, and maybe most important, the predictions of our model are unchanged whether the dealer and the arbitrageur's reaction times are respectively, say, 600 and 800 milliseconds on average ($\lambda = 1/600$ and $\gamma = 1/800$) or whether they are 100 times smaller. Indeed, this is traders' *relative* speed ($\frac{\lambda}{\gamma}$) that matters for dealers' exposure to the risk of trading at stale quotes with an arbitrageur (see eq.(4)), not their *absolute* speed. An increase in absolute speeds only means that the race between dealers and arbitrageurs is played over a shorter time interval.

As a robustness check, we have obtained data (from Thomson Reuters Tick History dataset) on trades and orders in Reuters D-3000 for the the three currency pairs in our sample over a more recent period (January 2, 2005-December 30, 2011). The data are identical to those used in our main tests, except that we do not observe the quantities posted at the best quotes. Hence, we cannot use the slope of the limit order book as a measure of illiquidity. Descriptive statistics for this alternative sample are provided in the Internet Appendix (see Table IA.4.1). Not surprisingly, the duration of arbitrage opportunities is smaller in the 2005-2011 sample (equal to 783 ms versus 1554 ms in the 2003-2004 period) because traders' absolute speeds have increased over time. More importantly, toxic arbitrage opportunities still occur frequently in recent years ($\varphi_t = 40.5\%$ in the 2005-2011 sample versus 50% in the 2003-2004 sample). Moreover, arbitrageurs' relative speed is higher in the 2005-2011 sample ($\pi_t^{tox} = 80\%$ versus 63.5% in 2003-2004).

In sum, despite the fact that trading is much faster in recent years, dealers are still exposed to the risk of trading at stale quotes with arbitrageurs. Thus, their bid-ask spreads should reflect this risk. To check this, we have estimated eq.(24) and eq.(26) for the 2005-2011 sample with OLS (without controlling for illiquidity on EBS as we do not have EBS data over the same period). Results are reported in Table IA.4.4 in the Internet Appendix.

We still observe a significant and positive relationship between the daily fraction of toxic arbitrage opportunities (φ_t) and illiquidity in the 2005-2011 sample (e.g., for the adverse selection component, $b_2 = 1.330$ with a t-stat of 7.06). Moreover, as predicted, arbitrage opportunities are shorter on days in which the fraction of arbitrage opportunities that are toxic is higher. The relationship between the likelihood that a toxic arbitrage terminates with a trade (π_t^{tox}) and illiquidity is positive but it is not statistically significant for some illiquidity measures. As previously observed, this might reflect the fact that arbitrageurs' relative speed is endogenous to illiquidity. We cannot address this issue because we have no instrument for π_t^{tox} over the 2005-2011 period . Interestingly, the effects of φ_t and π_t^{tox} on illiquidity are weaker during the crisis period (2008-2009), maybe because arbitrage was more difficult during the crisis.

4. Conclusions

At high frequency, transient demand shocks or delays in the adjustment of prices to news create very short lived arbitrage opportunities in pairs of assets with correlated payoffs. Arbitrage opportunities due to asyncronicities in the adjustment of prices to news are toxic because they expose dealers to the risk of trading with arbitrageurs at stale quotes. We show theoretically that, for this reason, more frequent toxic arbitrage opportunities and a faster arbitrageurs' response to these opportunities raise adverse selection costs for dealers and therefore impair liquidity. We provide evidence for these predictions using a sample of triangular arbitrage opportunities. Specifically we find that bid-ask spreads and adverse selection costs for the currency pairs in our sample are larger on days in which (a) the fraction of arbitrage opportunities that are toxic (i.e., due to news rather than transient demand shocks) is higher and (b) the frequency with which arbitrageurs successfully exploit these opportunities is higher.

One way to alleviate adverse selection costs due to toxic arbitrage opportunities is to reduce arbitrageurs' relative speed. For instance, as some market participants advocate, market orders could be put on hold for a very small random period of time before execution.[34] However, adverse selection costs of high frequency arbitrage must be balanced with its benefits, namely a greater price efficiency and a quicker provision of liquidity when arbitrage opportunities are due to transient demand shocks. Slowing down high frequency arbitrageurs is desirable only if associated adverse selection costs exceed these benefits. Our findings suggest that these costs can be significant. In contrast, little is known about the social value of accelerating by a few milliseconds the speed at which prices converge to efficient levels or at which arbitrageurs respond to price pressures. Future research should address this issue. It would help policy makers and market organizers to design trading rules that optimally balance the costs and benefits of high frequency arbitrage.

In any case, some care must be taken with our conclusions since our data cover only one type of the many high frequency arbitrage opportunities that are exploited by high speed arbi-

[34] See for instance "Interactive Brokers Group Proposal to Address High Frequency Trading" available at: "https://www.interactivebrokers.com/download/

trageurs. Our tests provide suggestive evidence for the mechanisms highlighted in our model but more research is needed to establish the robustness of our conclusions for other high frequency arbitrage opportunities (e.g., between ETFs and futures).

APPENDIX

Proof of Proposition 1

Using eq.(6), it is immediate that the unique bid-ask spread S_Y^* solving the zero profit condition (8) for market maker Y is $S_Y^* = 0$. Furthermore, solving eq.(5) for the bid-ask spread S_X^* such that $\Pi^X(S_X^*, \lambda^*, \gamma^*) = 0$ yields a unique solution for S_X^*, given by eq.(13) in Proposition 1. This proves that $S_Y^* = 0$ and S_X^* in eq.(13) are the unique zero profit bid-ask spreads.

Now consider the choices of their speeds by market maker X and the arbitrageur. The speed chosen by market maker X in equilibrium, λ^*, maximizes his expected profit, $\Pi^X(S_X^*, \lambda^*, \gamma^*)$. Hence, it must solve the following first order condition:

$$\frac{\partial \Pi^X(S_X^*, \lambda^*, \gamma^*)}{\partial \lambda} = (\frac{\varphi\alpha}{2})\frac{\gamma}{(\lambda+\gamma)^2}(\sigma - S_X^*) - \frac{c^m}{2} = 0. \tag{30}$$

This condition is necessary and sufficient because market maker X's expected profit is concave in its speed, λ. Similarly, the speed chosen by the arbitrageur in equilibrium, γ^*, maximizes her expected profit, $\Pi^a(S_X^*, S_Y^*, \lambda^*, \gamma^*) = 0$. Hence, it must solve the following first order condition:

$$\frac{\partial \Pi^A(S_X^*, 0, \lambda^*, \gamma^*)}{\partial \gamma} = (\frac{\varphi\alpha}{2})\frac{\lambda}{(\lambda+\gamma)^2}(\sigma - S_X^*) - \frac{c^a}{2} = 0, \tag{31}$$

where we have replaced S_Y^* by its equilibrium value, i.e., zero. Again, this condition is necessary and sufficient because the arbitrageur's expected profit is concave in its speed, γ.

Thus, we have a system of two equations (eq.(30) and eq.(31)) that must be satisfied by λ^* and γ^*. Solving this system of equations for λ^* and γ^*, we obtain a unique solution, which is as given in eq.(15) and eq.(16) in Proposition 1. Moreover, using the expressions for λ^* and γ^* in Proposition 1, one gets that $\frac{\gamma^*}{\lambda^*} = \frac{c^m}{c^a}$, which implies that $\pi^* = \frac{\gamma^*}{\lambda^*+\gamma^*} = \frac{\rho}{1+\rho}$.

Finally, observe that:

$$\Pi^X(S_X, \lambda, \gamma^*) < \Pi^X(S_X^*, \lambda, \gamma^*) < \Pi^X(S_X^*, \lambda^*, \gamma^*) \quad \forall S_X < S_X^*, \forall \lambda. \tag{32}$$

The first inequality follows from the fact that, other things equal, $\Pi^X(S_X, \lambda, \gamma^*)$ increases with S_X (see eq.(5)) and the second inequality follows from the fact that λ^* maximizes $\Pi^X(S_X^*, \lambda, \gamma^*)$ (Condition (11)). Now, as $\Pi^X(S_X^*, \lambda^*, \gamma^*) = 0$, we deduce from the previous equation that

$$\Pi^X(S_X, \lambda, \gamma^*) < \Pi^X(S_X^*, \lambda, \gamma^*) < 0 \quad \forall S_X < S_X^*, \forall \lambda. \tag{33}$$

This shows that the equilibrium $(S_X^*, S_Y^*, \lambda^*, \gamma^*)$ satisfies the no undercutting condition (10). Moreover, as S_X^*, S_Y^*, λ^*, γ^* are the unique solutions to eq.(8), (9), (11), and (12), we deduce that the competitive equilibrium of the model is unique.

Derivation of Implications 3 and 4

Using the fact that $\gamma^* = \rho\lambda^*$ in equilibrium, we deduce from eq.(17) that the expected duration of an arbitrage opportunity in equilibrium is:

$$\mathsf{E}(D) = \frac{(1+\rho) - \varphi}{\gamma^*(1+\rho)}. \tag{34}$$

Next, substituting the expression for the equilibrium spread, S_X^* (given in eq.(13)) in eq.(16), we obtain that:

$$\gamma^* = \frac{(\varphi\alpha\rho)(1-\alpha(2\varphi-1))}{(c^a(1+\rho)^2)(2\varphi\alpha\pi^*(\rho)(2-\pi^*(\rho)) + (1-\alpha(2\varphi-1)))}\sigma. \tag{35}$$

Substituting (35) in (34), we obtain:

$$\mathsf{E}(D) = \frac{h(\rho)(c^a(1-\varphi)+c^m)}{\varphi\alpha(1-\alpha(2\varphi-1))\sigma}, \tag{36}$$

with

$$h(\rho) = 2\varphi\alpha(\frac{2+\rho}{1+\rho}) + (1-\alpha(2\varphi-1))(\frac{2+\rho}{\rho}).$$

It is straightforward that $h(\rho)$ decreases with ρ. As ρ increases when c^a decreases, we deduce from eq.(36) that the expected duration of arbitrage opportunities decreases when c^a decreases (Implication 3).

Finally, using eq.(35), we also deduce that $\gamma^*(S_X^*)$ increases with φ if $\alpha(4\varphi-1) < 1$. We deduce from eq.(34) that the average duration of an arbitrage opportunity decreases with φ if $\alpha(4\varphi-1) < 1$ (Implication 4).

Triangular Arbitrage Opportunities: Example.

We illustrate our methodology to identify triangular arbitrage opportunities in our sample with an example. The table below gives best ask and bid prices for the three currency pairs in our sample at time t_0^-, just before an arbitrage opportunity.

Triangular Arbitrage Opportunities: An Example

Quotes	$B^{i/j}$	$A^{i/j}$
\$/€	1.074	1.0780
\$/£	1.6255	1.6265
£/€	0.6622	0.6632

These quotes are such that there is no triangular arbitrage opportunity. Now suppose that a sell limit order placed at 1.075 arrives in the USD/EUR market at time t_0. The new best quotes in this market become $A^{\$/€} = 1.075$ and $B^{\$/€} = 1.074$. If other rates are unchanged, we have $\hat{B}^{\$/€} = B^{\$/£} \times B^{£/€} = 1.0764$. As $\hat{B}^{\$/€} > A^{\$/€} = 1.075$, there is a triangular arbitrage opportunity.

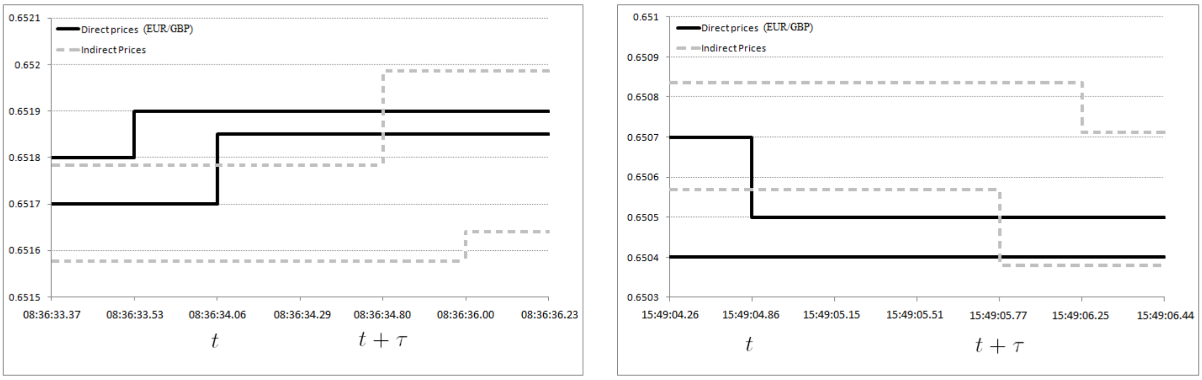
Now suppose that at date t_1, the limit sell order placed at $A^{\$/€} = 1.075$ is filled and that the new best ask price becomes again $A^{\$/€} = 1.078$. The arbitrage disappears and its duration is $t_1 - t_0$ seconds. We record that a trade terminates the arbitrage. Alternatively, suppose that at t_1, dealers in the USD/GBP market update their bid price at, say, $B^{\$/£} = 1.622$ so that $\hat{B}^{\$/€} = 1.074$. Thus, the arbitrage disappears but without any trade. In this case, we record that it terminates with a quote update.

Figures

Figure 1: Toxic versus Non-Toxic Arbitrage Opportunities

This figure illustrates how we classify triangular arbitrage opportunities into toxic and non-toxic opportunities. In each panel, the arbitrage opportunity starts at time t and ends at time $t+\tau$. The solid lines show the evolution of best ask and bid prices in the currency pair that initiates the arbitrage opportunity. The dashed lines show the evolution of best bid and ask synthetic quotes. In Panel A, we provide two examples of opportunities that we classify as toxic because they are associated with permanent shifts in exchange rates. In Panel B, we provide two examples of opportunities that we classify as non-toxic because the exchange rate in the currency pair initiating the arbitrage opportunity reverts to its level at the beginning of the opportunity.

Panel A: Toxic Arbitrage Opportunities



Panel B: Non-Toxic Arbitrage Opportunities

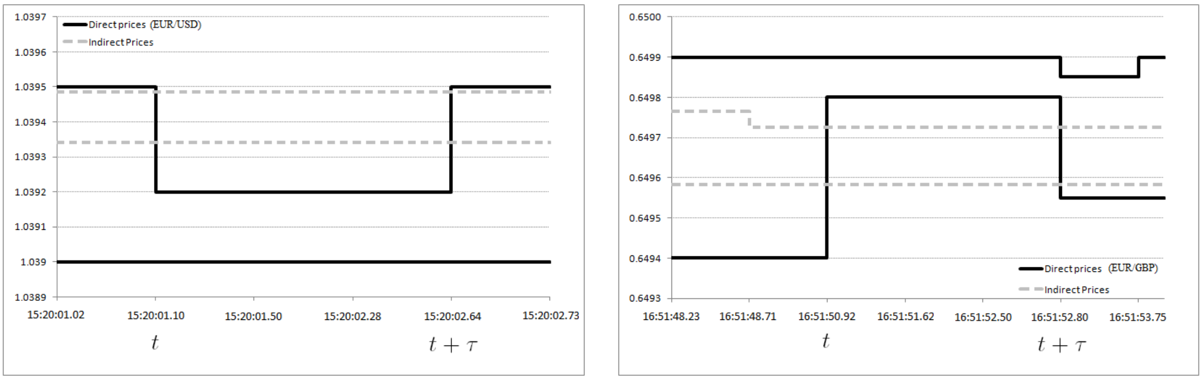
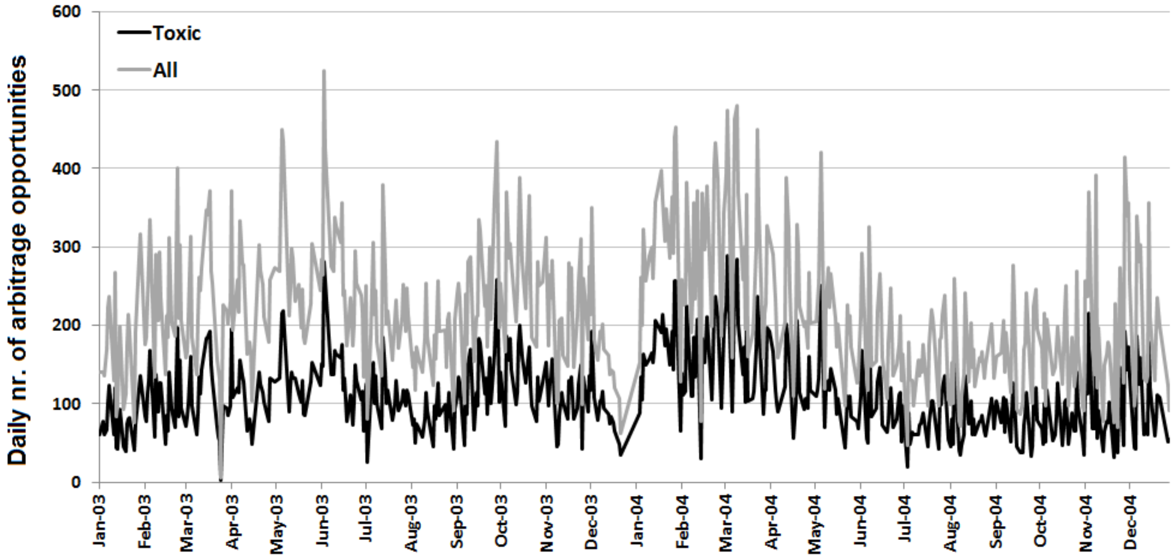


Figure 2: Number of Arbitrage Opportunities

Panel A shows the time series of the daily number of all triangular arbitrage opportunities (grey line) and toxic arbitrage opportunities (black line) in our sample. Panel B shows the average number of toxic and non-toxic arbitrage opportunities per hour in our sample. Time is GMT.

Panel A: Daily Numbers of Arbitrage Opportunities



Panel B: Intraday Pattern in the Number of Arbitrage Opportunities

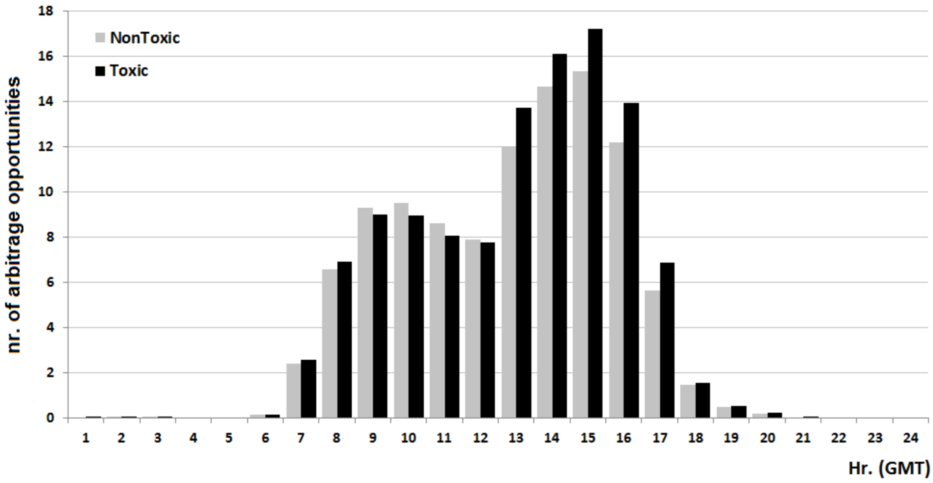
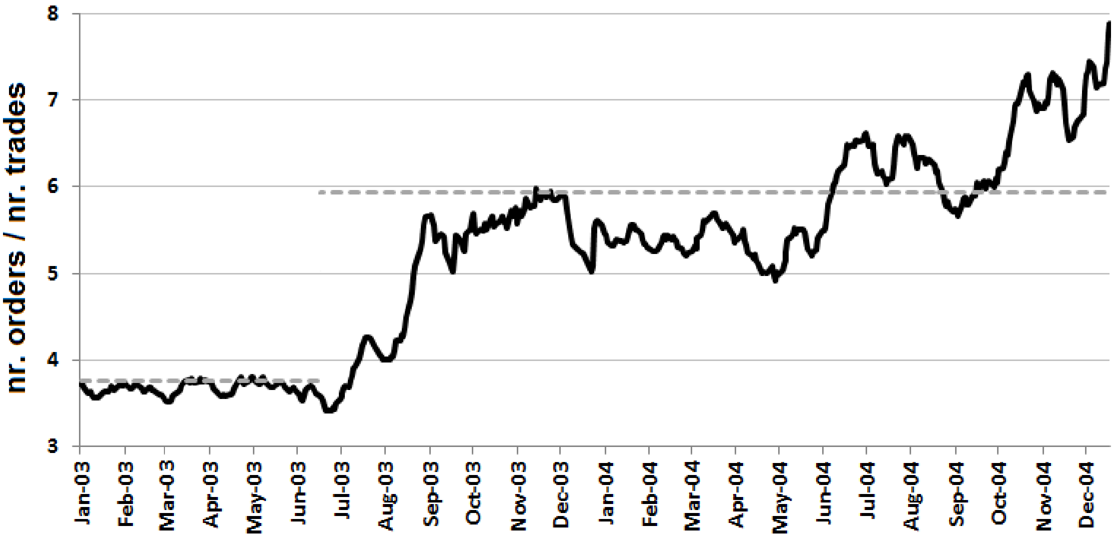


Figure 3: AutoQuote and Order-to-Trade Ratio.

This figure shows the evolution of the order-to-trade ratio (defined as the daily number of orders to the daily number of trades for the three currency pairs in our sample) from January 2003 to December 2004. The dashed lines indicate the average levels of the order to trade ratio before and after July, 1^{st} 2003.



Tables

Table 2: Descriptive Statistics

This table presents the descriptive statistics for the variables used in our tests for each currency pair $i \in \{GU, EU, EG)$, where indexes GU, EU, and EG refer to the GBP/USD, EUR/USD, and EUR/GBP currency pairs, respectively. Panels A provides descriptive statistics common to all opportunities while Panels B and C present descriptive statistics for variables that are specific to toxic (non-toxic) arbitrage opportunities. φ_t is the number of toxic arbitrage opportunities on day t divided by the number of arbitrage opportunities on this day; α_t is the number of all arbitrage opportunities on day t divided by the total number of trades on this day; $duration_t^{all}$ denotes the duration in seconds of all arbitrage opportunities on day t; $duration_t^{tox}$ ($duration_t^{nontox}$) denotes the average duration in seconds of toxic (non-toxic) arbitrage opportunities on day t; $nrarb_t^{tox}$ ($nrarb_t^{nontox}$) is the number of toxic (non-toxic) arbitrage opportunities on day t; π_t^{tox} (π_t^{nontox}) is the number of toxic (non-toxic) arbitrage opportunities on day t that terminate with a trade divided by the total number of toxic (non-toxic) arbitrage opportunities on this day; σ_t^{tox} (σ_t^{nontox}) is the average size of toxic (non-toxic) arbitrage opportunities on day t (in basis points); $profit_t^{tox}$ ($profit_t^{nontox}$) is the average profit in basis points on toxic (non-toxic) triangular arbitrage opportunities on day t (calculated as explained in Section 3.2). The sample period is from January 2, 2003 to December 30, 2004.

Panel A: The Arbitrage Mix

Variable	Mean	Std.Dev.	Min	Q1	Median	Q3	Max
$duration^{all}$	1.554	0.480	0.898	1.213	1.440	1.775	4.990
φ	0.501	0.060	0.280	0.463	0.505	0.545	0.650
α	0.065	0.016	0.029	0.052	0.064	0.077	0.109

Panel B: Toxic Arbitrage

$duration^{tox}$	1.710	0.524	0.858	1.344	1.615	1.960	4.715
$nrarb^{tox}$	112.0	49.62	2.000	76.00	105.0	137.0	288.0
π^{tox}	0.631	0.059	0.449	0.593	0.627	0.663	1.000
σ^{tox}	2.669	0.529	1.647	2.312	2.614	2.910	5.129
$profit^{tox}$	0.651	0.148	0.452	0.578	0.623	0.683	2.035

Panel C: Non-toxic Arbitrage

$duration^{nontox}$	1.448	0.505	0.842	1.101	1.335	1.671	6.768
$nrarb^{nontox}$	108.0	39.22	4.000	80.00	102.0	131.0	243.0
π^{nontox}	0.470	0.068	0.286	0.420	0.469	0.513	0.712
σ^{nontox}	2.641	0.634	1.653	2.270	2.539	2.871	10.41
$profit^{nontox}$	0.707	0.243	0.383	0.571	0.652	0.761	3.584

Table 3: Arbitrageurs' Relative Speed (π^{tox}), Arbitrage Mix (φ) and Liquidity

This table reports estimates of the following equation: $illiq_{it} = \omega_i + \xi_{t,m} + b_0 t + b_1 \pi_t^{tox} + b_2 \varphi_t + b_3 \alpha_t + b_4 \sigma_t^{tox} + b_5 vol_{it} + b_6 trsize_{it} + b_7 nrorders_{it} + b_8 ted_t + b_9 illiq_{it}^{EBS} + \varepsilon_{it}$, for $i \in \{GU, EU, EG\}$ where indexes GU, EU, and EG refer to the GBP/USD, EUR/USD, and EUR/GBP currency pairs, respectively. $illiq_{it}$ is one of our four proxies for illiquidity for currency i on day t: $spread_{it}$ is the average quoted bid-ask spread (in basis points) in currency pair i on day t; $espread_{it}$ (in basis points) is the average effective spreads in currency pair i on day t; $slope_{it}$ is the average slope of the limit order book in currency pair i on day t; $adv\_selection_{it}$, is the average 1-minute price impact of trades. Superscript EBS is used for measures of these variables computed using EBS data. φ_t is the number of toxic arbitrage opportunities on day t divided by the number of arbitrage opportunities on this day and π_t^{tox} is the number of toxic arbitrage opportunities on day t that terminate with a trade divided by the total number of toxic arbitrage opportunities on day t. We instrument π_t^{tox} with the introduction of AutoQuote on Reuters D-3000 (see the text). The first stage regression is: $\pi_t^{tox} = \omega_i + \xi_{t,m} + a_0 t + a_1 AD_t + a_2 \varphi_t + a_3 \alpha_t + a_4 \sigma_t^{tox} + a_5 vol_{it} + a_6 trsize_{it} + a_7 nrorders_{it} + a_8 ted_t + a_9 illiq_{it}^{EBS} + u_{it}$, where AD_t is a dummy variable equal to one after July 2003 and zero before. Other control variables are: α_t is the number of all arbitrage opportunities on day t divided by the total number of trades on this day; σ_t^{tox} is the average size of toxic arbitrage opportunities on day t (in basis points); vol_{it} is the realized volatility (in percentage) of 5-minutes returns for currency pair i on day t; $nrorders_{it}$ (in thousands) is the total number of orders (market, limit or cancelations) in currency pair i on day t; $trsize_{it}$ is the average daily trade size (in million) for currency pair i on day t; ted_t is the TED spread on day t, i.e., the difference between the LIBOR and the T-Bill rate on day t. In all regressions we include a currency pair fixed effect (ω_i), a monthly fixed effect ($\xi_{t,m}$) and a time trend (coefficients b_0 and a_0). t-statistics in parenthesis are calculated using robust standard errors correcting for heteroscedasticity and serial correlation. The sample period is from January 2, 2003 to December 30, 2004.

	spread		*espread*		*slope*		*adv_selection*	
	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage
AD	0.036 (5.37)		0.038 (5.61)		0.034 (5.12)		0.034 (5.22)	
π^{tox}		6.341 (4.55)		2.198 (4.00)		4.831 (4.68)		2.223 (3.43)
φ	-0.046 (-1.21)	0.875 (3.53)	-0.047 (-1.23)	0.393 (3.93)	-0.044 (-1.16)	0.575 (3.20)	-0.043 (-1.14)	0.428 (4.11)
α	-0.693 (-5.09)	8.243 (5.57)	-0.712 (-5.27)	3.551 (5.77)	-0.717 (-5.26)	6.015 (5.42)	-0.696 (-5.14)	4.144 (5.76)
σ^{tox}	-0.001 (-0.24)	0.121 (3.13)	-0.002 (-0.32)	0.119 (6.46)	-0.001 (-0.29)	0.038 (1.44)	-0.001 (-0.15)	0.059 (3.22)
vol	-0.008 (-0.45)	1.429 (7.47)	-0.007 (-0.40)	0.749 (8.11)	-0.004 (-0.22)	0.835 (6.41)	-0.010 (-0.58)	0.767 (7.82)
$trsize$	-0.018 (-0.52)	-0.109 (-0.46)	-0.023 (-0.68)	-0.054 (-0.60)	-0.019 (-0.56)	-0.229 (-1.33)	-0.019 (-0.57)	-0.138 (-1.45)
$nrorders$	-0.002 (-3.48)	-0.034 (-8.94)	-0.002 (-3.52)	-0.013 (-8.41)	-0.002 (-3.51)	-0.017 (-6.50)	-0.002 (-3.49)	-0.012 (-7.21)
ted	-0.111 (-2.21)	0.919 (2.23)	-0.114 (-2.27)	0.396 (2.45)	-0.129 (-2.55)	0.611 (1.98)	-0.108 (-2.15)	0.346 (2.01)
$illiq^{EBS}$	0.002 (1.82)	0.009 (1.00)	0.0042 (3.87)	-0.001 (-0.27)	0.0024 (1.48)	-0.003 (-0.39)	0.003 (2.87)	-0.006 (-1.39)
Adj. R^2	9.7%	70.4%	10.1%	86.3%	10.0%	39.2%	9.9%	74.0%
$F-stats$	28.8		31.4		26.2		27.2	
Obs.	1,449		1,449		1,449		1,449	
Currency FE	YES		YES		YES		YES	
Month FE	YES		YES		YES		YES	

Table 4: Duration of Arbitrage Opportunities

In this table, we present OLS estimates of the following equation: $\log(duration_t) = c_i + \xi_{t,m} + a_0 t + a_1 AD_t + a_2\varphi_t + a_3\alpha_t + a_4\sigma_t + a_5 vol_{it} + a_6 trsize_{it} + a_7 nrorders_{it} + a_8 ted_t + a_9 spread^{EBS}_{it} + u_{it}$ where $duration_t$ is the average duration of toxic arbitrage opportunities (*Toxic* column) on day t or all (both toxic and non-toxic) arbitrage opportunities (*All* column); AD (AutoQuote Dummy) is a dummy variable equal to one after July, 2003 and 0 before; φ_t is the number of toxic arbitrage opportunities divided by the number of arbitrage opportunities on day t; α_t is the number of all arbitrage opportunities on day t divided by the total number of trades on this day; σ^{tox}_t is the average size of arbitrage opportunities in day t (in basis points); vol_{it} is the realized volatility (in percentage) of 5-minutes returns for currency pair i on day t; $nrorders_{it}$ (in thousands) is the total number of orders (market, limit or cancelations) in currency pair i on day t; $trsize_{it}$ is the average daily trade size (in million) for currency pair i on day t; ted_t is the TED spread on day t, i.e., difference between the LIBOR and the risk-free T-Bill rate; $spread^{EBS}_{it}$ is the average quoted bid-ask spread (in basis points) in currency pair i on day t computed using the EBS data. In all regressions we include a currency pair fixed effect (ω_i), a monthly fixed effect ($\xi_{t,m}$, a dummy equal to one if day t is in month m) and a time trend (coefficients b_0 and a_0). t-statistics in parenthesis are calculated based on robust standard errors correcting for heteroscedasticity and serial correlation. The sample period is from January 2, 2003 to December 30, 2004.

	Toxic	All
AD	-0.067 (-3.68)	-0.068 (-4.34)
φ	-0.934 (-7.91)	-0.868 (-8.80)
α	-0.125 (-0.30)	-0.694 (-1.80)
σ	0.103 (6.38)	0.125 (8.76)
vol	-0.221 (-3.47)	-0.240 (-4.28)
$trsize$	0.159 (1.58)	0.103 (1.11)
$nrorders$	-0.023 (-15.8)	-0.024 (-17.8)
ted	-0.190 (-1.19)	-0.261 (-1.96)
$spread^{EBS}$	0.004 (1.01)	0.006 (1.53)
Adj. R^2	58.4%	67.3%
Currency FE	YES	YES
Month FE	YES	YES
Obs.	1,449	

Table 5: Toxic Arbitrage or Other Forms of Adverse Selection?

This table reports estimates of the following equation for $i \in \{GU, EU, EG\}$: $illiq_{it} = \omega_i + \xi_{t,m} + b_0 t + b_1 \pi_t^{tox} + b_2 \varphi_t + b_3 \alpha_t + b_4 \sigma_t^{tox} + b_5 vol_{it} + b_6 trsize_{it} + b_7 nrorders_{it} + b_8 ted + b_9 illiq_{it}^{EBS} + b_{10} D_t^{macro} + b_{11} std_t^{macro} + b_{12} vpin_{it} + b_{13} |oib|_{it} + \varepsilon_{it}, \quad for \quad i \in \{GU, EU, EG\}$. where indexes GU, EU, and EG refer to the GBP/USD, EUR/USD, and EUR/GBP currency pairs, respectively. $illiq_{it}$ is one of four proxies for illiquidity for currency i on day t: $spread_{it}$ is the average quoted bid-ask spread (in basis points) in currency pair i on day t; $espread_{it}$ (in basis points) is the average effective spreads in currency pair i on day t; $slope_{it}$ is the average slope of the limit order book in currency pair i on day t; $adv\_selection_{it}$, is the average 1-minute price impact of trades. Superscript EBS is used for measures of these variables computed using EBS data. φ_t is the number of toxic arbitrage opportunities on day t divided by the number of arbitrage opportunities on this day and π_t^{tox} is the number of toxic arbitrage opportunities on day t that terminate with a trade divided by the total number of toxic arbitrage opportunities on day t. We instrument π_t^{tox} with the introduction of AutoQuote on Reuters D-3000 (see the text). The first stage regression is: $\pi_t^{tox} = \omega_i + \xi_{t,m} + a_0 t + a_1 AD_t + a_2 \varphi_t + a_3 \alpha_t + a_4 \sigma_t^{tox} + a_5 vol_{it} + a_6 trsize_{it} + a_7 nrorders_{it} + a_8 ted_t + a_9 illiq_{it}^{EBS} + a_{10} D_t^{macro} + a_{11} std_t^{macro} + a_{12} vpin_{it} + a_{13} |oib|_{it} + u_{it}, \quad for \quad i \in \{GU, EU, EG\}$, where AD_t is a dummy variable equal to one after July 2003 and zero before. Other control variables are: α_t is the number of all arbitrage opportunities on day t divided by the total number of trades on this day; σ_t^{tox} is the average size of arbitrage opportunities in day t (in basis points); vol_{it} is the realized volatility (in percentage) of 5-minutes returns for currency pair i in day t; $trsize_{it}$ is the average daily trade size (in million) for currency pair i on day t; $nrorders_{it}$ (in thousands) is the total number of orders (market, limit or cancelations) in currency pair i on day t; D_t^{macro} is a dummy equal to one if there is at least one influential macro-announcement in the EMU, U.K., and U.S. areas on day t; std_t^{macro} is the average dispersion of professional forecasts on day t for influential macro-announcements (see the text); $|oib_t^i|$ is the absolute order imbalance in currency i on day t, defined as $|oib_t^i| = \frac{|buys_t^i - sells_t^i|}{buys_t^i + sells_t^i}$, where $buys^i$ ($sells^i$) is the number of buy (sell) market orders for currency pair i on day t excluding buy and sell market orders that terminate arbitrage opportunities; $vpin_{it}$ is a measure of adverse selection in currency pair i on day t (see Easley et al. (2012)); ted_t, is the TED spread on day t, i.e., the difference between the LIBOR and the T-Bill rate; In all regressions we include a currency pair fixed effect (ω_i), a monthly fixed effect ($\xi_{t,m}$) and a time trend (coefficients b_0 and a_0). t-statistics in parenthesis are calculated based on robust standard errors correcting for heteroscedasticity and serial correlation. The sample period is from January 2, 2003 to December 30, 2004.

Table 5 continued.

	spread		*espread*		*slope*		*adv_selection*	
	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage
AD	0.036 (5.36)		0.038 (5.59)		0.034 (5.13)		0.035 (5.23)	
π^{tox}		6.457 (4.61)		2.230 (4.06)		4.837 (4.72)		2.254 (3.53)
φ	-0.043 (-1.17)	0.845 (3.39)	-0.044 (-1.19)	0.386 (3.85)	-0.041 (-1.13)	0.563 (3.17)	-0.041 (-1.11)	0.419 (4.00)
α	-0.717 (-5.19)	8.378 (5.52)	-0.735 (-5.37)	3.586 (5.71)	-0.745 (-5.38)	6.084 (5.39)	-0.723 (-5.25)	4.180 (5.72)
σ^{tox}	-0.001 (-0.13)	0.119 (3.04)	-0.001 (-0.21)	0.118 (6.42)	-0.001 (-0.16)	0.036 (1.36)	-0.001 (-0.03)	0.059 (3.20)
vol	-0.007 (-0.37)	1.431 (7.30)	-0.006 (-0.34)	0.750 (7.95)	-0.003 (-0.17)	0.832 (6.30)	-0.009 (-0.50)	0.769 (7.65)
$trsize$	-0.007 (-0.24)	-0.211 (-0.90)	-0.012 (-0.39)	-0.078 (-0.86)	-0.008 (-0.26)	-0.263 (-1.58)	-0.009 (-0.29)	-0.175 (-1.85)
$nrorders$	-0.002 (-3.12)	-0.030 (-6.83)	-0.002 (-3.15)	-0.012 (-6.76)	-0.002 (-3.17)	-0.016 (-4.99)	-0.002 (-3.12)	-0.011 (-5.68)
ted	-0.119 (-2.33)	0.965 (2.28)	-0.121 (-2.38)	0.406 (2.46)	-0.138 (-2.71)	0.632 (2.01)	-0.115 (-2.28)	0.363 (2.06)
$illiq^{EBS}$	0.002 (1.75)	0.009 (1.02)	0.004 (3.84)	-0.001 (-0.28)	0.003 (1.56)	-0.004 (-0.42)	0.004 (2.87)	-0.006 (-1.40)
D^{macro}	-0.007 (-2.11)	0.046 (1.79)	-0.006 (-2.09)	0.012 (1.21)	-0.006 (-1.94)	0.026 (1.45)	-0.007 (-2.13)	0.009 (0.81)
std^{macro}	0.001 (1.79)	-0.004 (-0.83)	0.001 (1.79)	-0.001 (-0.45)	0.001 (1.82)	-0.003 (-0.96)	0.001 (1.88)	-0.001 (-0.43)
$vpin$	-0.101 (-1.25)	0.786 (1.78)	-0.102 (-1.26)	0.185 (1.06)	-0.106 (-1.31)	0.320 (0.93)	-0.100 (-1.24)	0.264 (1.39)
$\|oib\|$	0.062 (0.76)	0.031 (0.07)	0.057 (0.70)	-0.003 (-0.02)	0.073 (0.90)	-0.119 (-0.35)	0.061 (0.75)	0.048 (0.23)
Adj. R^2	10.6%	69.8%	10.9%	86.2%	10.9%	39.3%	10.8%	73.8%
$F-stats$	28.7		31.3		26.3		27.3	
Obs.	1,449		1,449		1,449		1,449	
Currency FE	YES		YES		YES		YES	
Month FE	YES		YES		YES		YES	

References

Adrian, T., and H. Shin. 2010. Liquidity and leverage. *Journal of Financial Intermediation* 19:418–437.

Akram, F., D. Rime, and L. Sarno. 2008. Arbitrage in the foreign exchange market: Turning on the microscope. *Journal of International Economics* 76:237–253.

Andersen, T. G., Bollerslev, T., F. X Diebold, and C. Vega. 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review* 93:38–62.

Bank for International Settlements. 2005. Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in April 2004 *Available at http://www.bis.org/publ/rpfx04.htm.*

Baron, M., J. Broogard, B. Hagstromer, and A. Kirilenko. 2016. Risk and return in high frequency trading. *Available at http://dx.doi.org/10.2139/ssrn.2433118.*

Ben-David, I., F. Franzoni, and R. Moussawi. 2012. ETFs, arbitrage, and shock propagation. *Working paper, Fisher College of Business.*

Biais, B., T. Foucault, and S. Moinas. 2015. Equilibrium fast trading. *Journal of Financial economics* 116:292–313.

Bjønnes, G., and D. Rime. 2005. Dealer behavior and trading systems in foreign exchange markets. *Journal of Financial Economics* 75:571–605.

Boulatov, A., T. Hendershott, and D. Livdan. 2013. Informed trading and portfolio returns. *Review of Economic Studies* 80:35–72.

Brogaard, J., B. Hagstromer, L. Norden, and R. Riordan. 2015. Trading fast and slow: Colocation and liquidity. *Review of Financial Studies* 28: 3407–3443.

Brogaard, J., T. Hendershott, and R. Riordan. 2014. High frequency trading and price discovery. *Review of Financial Studies* 27:2267–2306.

Brunnermeier, M. K., S. Nagel, and L. H. Pedersen. 2008. Carry trades and currency crashes. *Working Paper, NBER.*

Budish, E., E. Cramton, and J. Shim. 2015. The high frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics* 130:1547–1621.

Chaboud, A., B. Chiquoine, E. Hjalmarsson, and C. Vega. 2014. Rise of the machine: Algorithmic trading in the foreign exchange market. *Journal of Finance* 65:2045–2084.

Copeland, D., and T. Galai. 1983. Information effects on the bid-ask spread. *Journal of Finance* 38:1457–1469.

Duffie, D. 2010. Asset price dynamics with slow moving capital. *Journal of Finance* 65:1237–1267.

Duffie, D., N. Garleanu, and L. Pedersen. 2005. Over-the-counter markets. *Econometrica* 73:1815–1857.

Easley, D., R. Engle, M. O'Hara, and L. Wu. 2008. Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics* 6:171–207.

Easley, D., M. M. Lopez de Prado, and M. O'Hara. 2012. Flow toxicity and liquidity in a high frequency world. *Review of Financial Studies* 25:1457–1493.

Fenn, D., S. Howison, M. Mcdonald, S. Williams, and N. Johnson. 2009. The mirage of triangular arbitrage in the spot foreign exchange market. *International Journal of Theoretical and Applied Finance*, 12:1105–1123.

Fong, W., G. Valente, and J. Fung. 2008. Market liquidity and FX arbitrage in emerging markets: Evidence from Hong Kong tick data. *Working Paper, Hong Kong Institute for Monetary Research.*

Foucault, T., J. Hombert, and J. Rosu. 2016. News trading and speed. *Journal of Finance* 61:335–382.

Foucault, T., A. Röell, and P. Sandås. 2003. Market making with costly monitoring: An analysis of the SOES controversy. *Review of Financial Studies* 16:345–384.

Gagnon, L., and G. Karolyi. 2009. Information, trading volume, and international stock returns comovements: Evidence from cross-listed stocks. *Journal of Financial and Quantitative Analysis* 44:953–986.

Gagnon, L., and G. Karolyi. 2010. Multi-market trading and arbitrage. *Journal of Financial Economics* 53:53–80.

Garvey, R., and F. Wu. 2010. Speed, distance, and electronic trading: New evidence on why location matters. *Journal of Financial Markets* 13:367–396.

Green, T. C. 2004. Economic news and the impact of trading on bond prices. *Journal of Finance* 59:1201–1234.

Gromb, D., and D. Vayanos. 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics* 66:361–407.

Gromb, D., and D. Vayanos. 2010. Limits to arbitrage: The state of the theory. *Annual Review of Financial Economics* 98:251–275.

Grossman, S., and M. Miller. 1988. Liquidity and market structure. *Journal of Finance*. 43:617–637.

Hendershott, T., C. M. Jones, and A. J. Menkveld. 2011. Does algorithmic trading improve liquidity? *Journal of Finance* 66:1–34.

Hendershott, T., and A. Menkveld. 2014. Price pressures. *Journal of Financial Economics* 114:405–423.

Hendershott, T., and P. Moulton. 2011. Automation, speed, and stock market quality: The NYSE's hybrid. *Journal of Financial Markets* 14:568–604.

Ho, T., and H. R. Stoll. 1981. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9:47–73.

Hoffmann, P. 2014. A dynamic limit order market with fast and slow traders. *Journal of Financial Economics* 113:159–169.

Holden, C. 1995. Index arbitrage as cross sectional market making. *Journal of Futures Markets* 15:423–455.

Holden, C., and S. Jacobsen. 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *Journal of Finance* 69:1747–1785.

Hu, X., J. Pan, and J. Wang. 2013. Noise as information for illiquidity. *Journal of Finance* 68:2341–2382.

Ito, T., K. Yamada, M. Takayasu, and H. Takayasu. 2013. Free lunch! Arbitrage opportunities in the foreign exchange markets. *Working Paper, NBER.*

Jovanovic, B., and A. Menkveld. 2012. Middlemen in limit order markets. *Working Paper, University of Amsterdam.*

King, M. R., C. Osler, and D. Rime. 2012. *Foreign exchange market structure, players and evolution.* Handbook of Exchange Rates. Wiley.

King, M. R., and D. Rime. 2010. The $4 trillion question: What explains FX growth since the 2007 survey? *BIS Quarterly Review* 27–40.

Kozhan, R., and W. W. Tham. 2012. Execution risk in high-frequency arbitrage. *Management Science* 58:2131–2149.

Krishnamurthy, A. 2002. The bond/old bond spread. *Journal of Financial Economics* 66:463–506.

Kumar, P., and D. Seppi. 1994. Information and index arbitrage. *Journal of Business* 67:481–509.

Mancini-Griffoli, T., and A. Ranaldo. 2011. Limits to arbitrage during the crisis: funding liquidity constraints and covered interest parity. *Working paper, Swiss National Bank.*

Marshall, B., S. Treepongkaruna, and M. Young. 2008. Exploitable arbitrage opportunities exist in the foreign exchange market. *mimeo.*

Menkveld, A., and M. Zoican. 2014. Need for speed? Exchange latency and liquidity. *Working Paper, University of Amsterdam.*

Naik, N., and P. Yadav. 2003. Do dealer firms manage inventory on a stock-by-stock or portfolio basis? *Journal of Financial Economics* 69:325–353.

Osler, C. 2008. *Foreign exchange microstructure: A survey.* Springer Encyclopedia of Complexity and System Science. Springer.

Pasquariello, P. 2014. Financial market dislocations. *Review of Financial Studies* 27:1868–1914.

Pasquariello, P. 2016. Government intervention and arbitrage. *Working Paper, University of Michigan.*

Pasquariello, P., and C. Vega. 2007. Informed and strategic order flow in the bond markets. *Review of Financial Studies* 20:1975–2019.

Pierron, A. 2007. Electronic platforms in foreign exchange trading. Discussion paper. Celent. Paris.

Rime, D., L. Sarno, and E. Sojli. 2010. Exchange rate forecasting, order flow and macroeconomic information. *Journal of International Economics* 80:72–88.

Roll, R., E. Schwartz, and A. Subrahmanyam. 2007. Liquidity and the Law of One Price: The case of future-cash basis. *Journal of Finance* 62:2201–2234.

Rosch, D. 2014. The impact of arbitrage on market liquidity. *Working paper, Erasmus University.*

Schultz P., and S. Shive 2010. Mispricing of dual class shares: Profit opportunities, arbitrage, and trading. *Journal of Financial Economics* 98:524–549.

Stoll, H. 2000. Friction. *Journal of Finance* 55:1479–1514.

U.S. Securities and Exchange Commission. 2010. U.S. concept release on equity market structure. *17 CFR Part 242, Release No. 34-61358*, File No. S7–02–10.