Liquidity Provision in the Foreign Exchange Market *

Florent Gallien[a], Serge Kassibrakis[a], Nataliya Klimenko[a], Semyon Malamud[b,c] and Alberto Teguia[d]

[a]Swissquote

[b]Swiss Finance Institute

[c]Ecole Polytechnique Fédérale de Lausanne

[d]Sauder School of Business, University of British Columbia

This version: October 17, 2018

Abstract

Foreign exchange operates as a two-tiered over-the-counter (OTC) market dominated by large, strategic dealers. Using proprietary high frequency data on quotes by the largest foreign exchange dealer banks in the Dealer-to-Client market, we find a significant heterogeneity in their behavior. We develop a model of strategic competition that accounts for this heterogeneity and the two-tier market structure and allows us to link prices and bid-ask spreads in the D2C and D2D market segments. We use the model to recover dealers' risk aversions and inventories from their quotes in the D2C segment and construct an endogenous measure of systemic, non-diversifiable risk capturing the cross-sectional liquidity-risk mismatch. Consistent with the model predictions, we find that liquidity mismatch negatively predicts prices in the D2D market.

Keywords: Liquidity, Foreign Exchange, OTC markets, Price Impact, Market Power

JEL Classification Numbers: F31, G12, G14, G21

*We thank Alain Chaboud, Alexander Michaelides, Andreas Schrimpf, and Adrien Verdelhan, as well as conference participants at the 3rd Annual CEPR Symposium on Financial Economics for helpful comments and remarks.

1 Introduction

The microstructure of the foreign exchange market is extremely complex. Trading is decentralized and has a pronounced two-tier structure, characterized by two important segments: the dealer to customer (D2C) and the dealer to dealer (D2D) segments. Importantly, liquidity provision in the D2C segment is fragmented, dominated by a small number of large dealer banks,[1] and these banks are highly heterogeneous in their risk bearing capacity. As such, dealer banks follow distinct strategies in their market making and liquidity provision. This heterogeneity combined with market fragmentation implies that the distribution of risk exposures (inventories) across dealers matters for pricing and market liquidity. The goal of this paper is to develop a model that accounts for these realistic features of foreign exchange markets and then test this model using data on dealer quotes and spreads in the D2C and D2D segments.

Our model features multiple, strategic dealers that compete in providing liquidity to their customers, absorb (part of) customers inventory shocks, and then trade with each other in the inter-dealer market to share these inventories. Importantly, dealers (liquidity providers, henceforth LPs) are heterogeneous in their inventory holding costs (henceforth, risk aversions), determining their willingness to provide liquidity and hence the spreads they offer to customers in the D2C segment. A customer's problem reduces to optimally splitting the order across multiple dealers to minimize trading costs. By contrast, an LP's problem is much more complex due to the non-trivial tradeoff induced by the two-tier market structure. First, dealer's decision to set the mid-price and the bid-ask spread in the D2C market is determined by his inventory management needs and his competition for order flow with other dealers. Second, the less a given dealer provides liquidity in the D2C market, the more will customers trade with other dealers, and hence the more will be the size of the inventory these other dealers will be trying to offload in the D2D market. Most importantly,

[1] According to recent BIS statistics (see Moore et al. (2016)), in most of the top trading jurisdictions, less than ten dealers account for 75% of FX turnover.

since dealers are large and strategic, the D2D market is not perfectly liquid, and dealers have (endogenous) price impacts that are also heterogeneous: Indeed, less risk averse dealers face a less elastic residual supply curve because they are trading with high risk aversion dealers who are reluctant to provide liquidity. As a result, less risk averse LPs have a higher price impact. This heterogeneity in price impacts combined with the D2D market clearing leads to an illiquidity-adjusted version of the Capital Asset Pricing Model (CAPM), whereby equilibrium D2D price is given by the fundamental value net of the risk premium, given by an *endogenous aggregate risk portfolio.* In stark contrast to the classical CAPM, in the presence of imperfect competition the aggregate risk portfolio is different from the market portfolio and depends on the distribution of inventories across LPs. Since less risk averse LPs have higher price impact, they end up trading less and hence their inventory gets a lower weight in the risk premium. The difference between non-diversifiable risk (risk tolerance-weighted aggregate inventory) and the market portfolio (equal-weighted aggregate inventory) is what we call the "liquidity mis-match". By definition, this liquidity mis-match is given by a (weighted) difference between inventories of low risk aversion LPs and those of high risk aversion LPs. This mis-match influences equilibrium allocations of risk across LPs and hence affects the market clearing price in the D2D market. Furthermore, the cross-sectional dispersion in risk aversions impacts the joint ability of dealers to provide liquidity in the D2D market and hence influences the D2D bid-ask spreads. As a result, our model generates two main predictions linking dealers' heterogeneity to quotes and spreads in the two market segments:

(1) Liquidity mis-match negatively predicts D2D prices;

(2) Cross-sectional dispersion in risk aversions negatively predicts spreads in the D2D market.

(3) Customers' net order flow negatively predicts D2D prices.

The intuition behind (1) is straightforward: High risk version LPs absorb less shocks in the D2C market and participate less in trading in the D2D market; hence, their inventories get lower weights in the endogenous aggregate risk portfolio. Item (2) is more subtle and is driven by a convexity effect: Dealers' liquidity provision (as captured by the slope of their demand schedule in the D2D market) is inversely proportional to their risk aversion, and total D2D market liquidity is given by the total slope of dealers' demand. Since the function $1/x$ is convex, total slope increases in the dispersion of risk aversions. Finally, item (3) is a standard CAPM effect that does not depend on dealer heterogeneity: When customers offload their supply shock on dealers, total dealer inventory increases; hence, so does the total supply in the D2D market, pushing prices down.

We test predictions (1)-(3) using proprietary high frequency data on quotes in the FX spot market for the EURUSD currency pair by the major dealers in the dealer-to-client market segment, provided by a major Swiss retail aggregator. We merge these data with data on quotes and spreads for the same currency pair from Electronic Broking Services (EBS), one of the largest inter-dealer FX platform in the world.[2] Retail aggregators (RAs) act as intermediaries between retail clients and large foreign exchange dealers (LPs). These LPs compete for RAs order flow by providing high frequency quotes (bids and asks). Observing these quotes, RAs decide endogenously how much of the retail clients' order flow to keep, and optimally split the rest of their inventory shocks across different LPs. In our model, the vector of levels (mid prices) and the vector of bid-ask spreads of different LPs are both functions of their inventories and risk aversions. We derive these functions analytically and then invert them to recover LPs inventories and risk aversions. Given these inventories and risk aversions, we compute market clearing prices in the D2D market that come out of (imperfect) inventory sharing between the (large strategic) dealers. Importantly, since the retail order flow of the RA comes almost exclusively from Switzerland, and we only consider

[2]According to Mancini et al. (2013), EBS is the largest inter-dealer FX trading platform, accounting for 60% of total D2D volume.

the EURUSD currency pair, it is safe to assume that our RA does not have any informational advantage relative to the LPs, who are all large international investment banks. Thus, we believe that our assumption of pure inventory-driven trades is well justified.

The data strongly support the key predictions (1)-(2) above. In particular, liquidity mis-match is non-trivial, and has a significant impact on D2D prices. We then use the RA customers' net order flow as a (noisy) proxy for the global retail order flow to test (3) and also find strong support for this prediction. The fact the order flow of Swiss retail clients predicts D2D trades in the EURUSD currency pair (the most liquid FX pair in the world) is surprising.

2 Literature Review

Motivated by the failure of macroeconomic models to explain exchange rate dynamics,[3] a growing set of papers shows how order flow and dealer inventories serve as important determinants of exchange rates.

Lyons (1995) is one of the first to provide strong empirical evidence that dealers' actively control their inventories and studies how this inventory control creates a link between order flow and exchange rates. In particular, consistent with classical theories (see, e.g., Ho and Stoll (1981)), Lyons (1995) shows that LPs "shade prices": That is, they shift prices in the direction opposite to their inventory. While such price shading has been documented in other markets,[4] recent studies find no evidence of price shading in the FX markets. See, for example, Bjønnes and Rime (2005) and Osler et al. (2011). In contrast to these papers, and in agreement with Lyons (1995), our empirical results provide strong evidence for price shading.[5]

[3]The fact that exchange rates are only weakly related to macroeconomic fundamentals is known as Meese and Rogoff (1983) exchange rate disconnect puzzle.

[4]See, e.g., (Madhavan and Smidt, 1993) and (Dunne et al., 2010).

[5]In particular, in our additional (non-reported) analysis, we find that a large trade by the retail aggregator always leads all dealers to adjust the price consistent with the shading theory. As we explain above, the

In a seminal contribution, Evans and Lyons (2002b) develop the first theoretical model that accounts for the two-tier structure of the FX market and derive an endogenous link between order flow and exchange rates. Consistent with the constraints that the real world FX dealers face, Evans and Lyons (2002b) assume that dealers need to hold zero inventory overnight. As a result, dealers optimally shade their prices to achieve the zero inventory target. It is this price shading that leads to a contemporaneous relation between order flow and exchange rates. Similarly to Evans and Lyons (2002b), ours is a pure inventory-theoretic model.[6] However, it differs from that of Evans and Lyons (2002b) in several important dimensions. First, we assume that dealers are heterogeneous in their risk bearing capacities. Second, we introduce strategic competition between dealers for order flow in the D2C market. Third, we assume that dealers are also strategic when trading in the D2D market. This assumption is crucial for our main results: It implies that dealers have price impact and hence the joint distribution of inventories and price impacts (as captured by the liquidity mis-match) matters for equilibrium prices and allocations. Finally, we apply our model to very short horizons (up to ten seconds), and hence we do not need to impose the zero inventory constraint. We believe that all these new ingredients are important for real world markets and allow us to capture new effects that have not been studied in the literature before.

Evans and Lyons (2002b) test their model predictions empirically and provide a strong evidence for a contemporaneous link between order flow and exchange rates using data on Deutsche Mark/US dollar exchange rates. Their empirical findings have been further confirmed in numerous papers using various data sets. See, for example, Evans and Lyons (2002a), Hau et al. (2002), Fan and Lyons (2003), Froot and Ramadorai (2005), Bjønnes et al.

nature of our data suggests that LPs do not learn much from the RA order flow and hence the shading behaviour is most probably not driven by asymmetric information.

[6]See, Bacchetta and Van Wincoop (2006), Evans et al. (2011), Evans and Lyons (2005c); Cao et al. (2006), Frankel et al. (2009), Lyons et al. (2001), King et al. (2010), Michaelides et al. (2018), and Gargano et al. (2018) for FX microstructure models of exchange rates that rely on private, heterogeneous information. Incorporating such informational asymmetries into our model is an important direction for future research.

(2005), Danielsson and Love (2006), Killeen et al. (2006), Berger et al. (2008), Brunnermeier et al. (2008), King et al. (2010), Rime et al. (2010), Breedon and Vitale (2010), and Bjønnes et al. (2011). Going beyond the contemporaneous relationships, Evans and Lyons (2005b), Danielsson et al. (2012), and Evans and Rime (2016) find empirical evidence that order flow can be used to forecast exchange rate returns at horizons ranging from one day to one month. In this paper, our focus in on such predictive relationships, but at mush shorter horizons. We use our model to derive them explicitly and link them to risk aversion heterogeneity and liquidity mismatch, a new endogenous object that is unique to our model.

Our main predictions derive from dealers' heterogeneity. The presence of such heterogeneity is strongly supported by data. For example, Evans (2002) finds that most of the short-term volatility in exchange rates is due to heterogeneous trading decisions by dealers; similarly, Bjønnes and Rime (2005) document significant differences in dealers' trading styles, especially related to how they actually control their inventories. Our data on dealer quotes in the D2C market also suggests the presence of significant heterogeneity: Both the bid-ask spreads and the sensitivity of quotes to shocks are highly heterogeneous across dealers.

Most of our key predictions depend crucially on the fact that the FX market is not perfectly liquid and dealers have price impact. While the idea that price impact is linked to order flow is not new (see, for example, Evans and Lyons (2005a)), to the best of our knowledge our model is the first to micro-found this price impact in a model that accounts for the two-tier structure of the FX market. In particular, our model can be used to actually recover market liquidity from LP quotes, providing an explicit, micro-founded measure of liquidity risk and shedding new light on the important findings of Banti et al. (2012) and Mancini et al. (2013).

Finally, we also note that while our focus in this paper is on exchange rates, our theoretical model is applicable to any market with a pronounced two-tier structure, such as, e.g., the bond market (see, Brandt and Kavajecz (2004)) and the CDS market (see, Collin-

Dufresne et al. (2016)). Furthermore, while we take the agents' demand shocks as given, developing a theory of exchange rates that incorporates dealers (intermediaries) and (other) market microstructure frictions into a macroeconomic model would be a natural extension of our analysis. Recent papers show that market fragmentation (two tier market structure) combined with either limited dealers' risk bearing capacity (Gabaix and Maggiori (2015)) or dealers' market power (Malamud and Schrimpf (2017)) helps explaining several known puzzles about exchange rates behaviour in the context of a macro-economic model. We believe that future research combining the ideas from these papers with a fully micro-founded imperfect competition intermediation model in our paper may lead to further important insights.

3 The Model

There are five time periods, $t = -1,\ 0-\ ,\ 0,\ 1,\ 2$ and two tradable assets, a risk free asset with rate of return normalized to zero, and a risky asset with a random payoff d at time $t = 2$. We assume that d is normally distributed with mean $\bar{d}$ and variance σ_d^2. The market is populated by M heterogeneous dealers (liquidity providers), indexed by $i = 1, \cdots, M$, and n ex-ante identical customers (liquidity consumers), indexed by $j = 1, \cdots, n$. All agents start with zero asset inventories. Then, the time line is as follows

- at time $t = -1$, dealers' inventory shocks $x_i,\ i = 1, \cdots, M$ are realized and are public information.[7]
- At time $t = 0-$, customers' inventory shocks vector $\Theta = \{\theta_j\}_{j=1}^n$ is realized; each shock is customers' private information. We assume that these shocks are independent and

[7]This assumption is made for simplicity. This private information is irrelevant for the D2D trading round because of the ex-post nature of the D2D market mechanism. As for the D2C round, uncertainty about other dealers' holdings would slightly complicate the analysis, but would not alter our main results. It would however become quite important in a multi-period model, in which case it would introduce signaling and adverse selection effects. As our goal is to focus on risk sharing, we abstract from these effects.

identically distributed across customers, and are drawn from a binary distribution: θ takes two values, $\theta_H, \theta_L \in \mathbb{R}$. Parameters $\bar{\theta} = E[\theta]$ and $\sigma_\theta^2 = \mathrm{Var}[\theta]$ are public information.

- At time $t = 0$, dealers trade in the dealer-to-customer (D2C) market using a request-for-quote protocol. In this protocal, each dealer submits a (dealer-specific) binding price schedule $p_i(q),\ i = 1, \cdots, M$ to each customer, describing the per-unit price at which he is willing to sell a lot of q units. Given the quoted price schedules, each customer j optimally chooses the vector of quantities $q_j = (q_{j,i})_{i=1}^M$ to trade with the dealers, and then pays the total price

$$\pi(q_j) \ = \ \sum_{i=1}^{M} q_{j,i}\, p_i(q_{j,i})$$

to the dealers, ending up holding a total of

$$\bar{q}_j \ = \ \theta_j \ + \ \sum_{i=1}^{M} q_{j,i}$$

units of the asset. After this trading round, dealer i receives

$$\Pi_i(Q_i) \ \equiv \ \sum_{j=1}^{n} p_i(q_{j,i}) q_{j,i} \tag{1}$$

in cash from the customers, and ends up holding

$$\chi_i \ = \ x_i \ - \ \sum_{j=1}^{n} q_{j,i} \tag{2}$$

units of the asset.

- At time $t = 1$, dealers trade in the centralized inter-dealer market to rebalance their inventories. We assume that this market operates as the standard uniform-price double auction (see, e.g., Kyle (1989); Vives (2011); Rostek and Weretka (2015); Malamud and Rostek (2017)). Dealer i submits a (net) demand schedule $Q_i(\mathcal{P}^{D2D}) : \mathbb{R} \rightarrow R$, which specifies demanded quantity of the asset given its price $\mathcal{P}^{D2D}$ in the inter-dealer market. All dealers are strategic; in particular, there are no noise traders. As is standard in strategic centralized market models for divisible goods or assets, we study the Nash equilibrium in linear bid schedules (hereafter, *equilibrium*). With divisible goods, equilibrium is invariant to the distribution of independent private uncertainty.[8] We denote by $Q_i(\chi_i, \mathcal{P}^{D2D})$ the net trade of dealer i with inventory (2) in the D2D market, so that, post-D2D trade, the dealer ends up with an inventory of

$$\tilde{\chi}_i \;=\; \chi_i \;+\; Q_i(\chi_i, \mathcal{P}^{D2D}) \tag{3}$$

- At time $t = 2$, the asset pays off.

We assume that agents incur quadratic costs for holding inventories, which is equivalent to linearly decreasing marginal values. Importantly, these inventory holding costs are heterogeneous across dealers: while customers are assumed to be homogeneous and all have the same cost γ, dealers are heterogeneous, with dealer i incurring the cost Γ_i.[9] Thus, customers'

[8] That is, the linear Bayesian Nash Equilibrium with independent private endowment values has an *ex post* property and coincides with the linear equilibrium that is robust to adding noise in trade (robust Nash Equilibrium; e.g., Vayanos (1999); Rostek and Weretka (2015)). Equilibrium is *linear* if schedules have the functional form of $q_i(\cdot) = \alpha_0 + \alpha_{i,q} q_i^0 + \alpha_{i,p} p$. Strategies are not restricted to linear schedules; rather, it is optimal for a trader to submit a linear demand given that others do. The approach of analyzing the symmetric linear equilibrium is common in centralized market models (e.g., Kyle (1989), Vayanos (1999), Vives (2011)). Our analysis does not assume equilibrium symmetry. Since equilibrium schedules are optimal even if traders learn the independent value endowments q_i^0 (or equivalently, stochastic marginal utility intercepts, $\tilde{d} = d - \alpha \Sigma \tilde{q}_i^0$) of all other agents, equilibrium is *ex post* Bayesian Nash. The key to the *ex post* property is that permitting pointwise optimization – for each price – equilibrium demand schedules are optimal for any distribution of independent private information and are independent of agents' expectations about others' endowments.

[9] The assumption of homogeneous customers is without loss of generality. By contrast, allowing for dealers' heterogeneity is crucial for matching their highly heterogeneous empirically observed behavior.

total expected utility is given by

$$u(q_j) \;=\; -\pi(q_j) \;+\; E\left[d\cdot\bar{q}_j \;-\; 0.5\gamma\,\bar{q}_j^2\right]\,, \tag{4}$$

while dealers' total expected utility is given by

$$U_i \;=\; E\left[\Pi_i(Q_i) \;+\; \mathcal{U}_i(\tilde{\chi}_i, \mathcal{P}^{D2D})\right]\,, \tag{5}$$

where $\Pi_i(Q_i)$ is the total transfer (1) received from customers, while

$$\mathcal{U}_i(\tilde{\chi}_i, \mathcal{P}^{D2D}) \;=\; d\cdot\tilde{\chi}_i - 0.5\Gamma_i\tilde{\chi}_i^2 - \mathcal{P}^{D2D}\,Q_i(\chi_i, \mathcal{P}^{D2D}) \tag{6}$$

is dealers' quadratic utility of his post-D2D trade inventory (3) net of the price $\mathcal{P}^{D2D}Q_i(\chi_i, \mathcal{P}^{D2D})$ paid for buying $Q_i(\chi_i, \mathcal{P}^{D2D})$ units of the asset in the D2D market at the price $\mathcal{P}^{D2D}$.

The strategic aspect of the interaction between the dealers make the model inherently complex. Each dealer selects a whole mechanism (the price schedule $p_i(q)$), while at the same time competing with other dealers' mechanisms and, simultaneously, taking into account the fact that other dealers effectively serve as liquidity providers to each other during the second stage of the game (the inter-dealer trading). Such games in competing mechanisms are known to be extremely complex; even with homogeneous traders, optimal price schedules are highly non-linear and a symmetric equilibrium often fails to exist. See, e.g., Biais et al. (2000, 2013), and Back and Baruch (2013). In our paper, the problem is much more involved because dealers are asymmetric and there is a second stage rebalancing, introducing another dimension to the strategic interaction. However, the assumption that customer types θ are binary implies that the dealers' screening problem is simple: Without loss of generality, we may assume that dealers always use linear demand schedules: $p_i(q) \;=\; \alpha_i \;+$

$\lambda_i^{-1} q$ with some $\alpha_i, \lambda_i^{-1} \in \mathbb{R}\,,\ ,\ i = 1, \cdots, M$.[10] The linearity of the schedule significantly simplifies the analysis, while allowing us to capture two most important dimensions of the price schedule: the level effect (through the α_i parameter) and the liquidity effect (through the slope parameter λ_i, that effectively measures the amount of liquidity that dealer i is providing to customers; $2\lambda_i^{-1}$ is the bid-ask spread for trading one unit of the asset). Thus, under the binary demand uncertainty in the D2C market, we can follow the standard route used in most of the market microstructure literature and confine our attention to linear equilibria, characterized in the following definition.

Definition 1 *A linear Nash equilibrium is a collection of policies:*

- *price schedules* $p_i(q) \;=\; \alpha_i \;+\; \lambda_i^{-1} q$ *in the D2C market segment,* $i = 1, \cdots, M$;
- *customer demand*

$$q_j(\theta) = (q_{j,i}(\theta))_{i=1}^M,\ q_{j,i}(\theta) \;=\; q_i^{(0)} + q_i^{(1)} \theta_j\,;$$

- *dealer demand schedules* $Q_i(\mathcal{P}^{D2D}) \;=\; Q_i^{(0)} + Q_i^{(1)} \mathcal{P}^{D2D}$ *in the inter-dealer market ;*

such that

- *dealer demand schedules form a robust Nash equilibrium in the D2D market (that is, each schedule* $Q_i(\mathcal{P}^{D2D})$ *is the best response to* $\{Q_j(\mathcal{P}^{D2D})\}_{j\neq i}$ *among all possible (linear and non-linear) demand schedules;*
- *dealer price schedules in the D2C market maximize dealers' expected utility* (5) *among all possible price schedules, given customers' demand functions* $q(\theta)$ *and given the equilibrium allocation from the second stage game;*
- *customers' demand maximizes customers' utility* (4).

[10]Indeed, dealers rationally anticipate that customer demand can take only two values in equilibrium, depending on the realization of $\theta = \theta_H,\ \theta_L$.

Given the definition of an equilibrium, we follow the standard backward induction procedure: First, we solve for the unique robust linear Nash equilibrium of the second stage game (the D2D market). Second, we use this equilibrium to calculate dealers' utilities (5). Third, we solve for customers' optimal demand given the linear dealer price schedules. Fourth, we use customers demand schedules as well as the dealers utilities from the second trading round to solve for the equilibrium in the liquidity provision game in the D2C market, whereby each dealer optimal chooses his linear price schedule taking as given the price schedules of other dealers.

4 Equilibrium in the D2D market

In this section, we use the results in Malamud and Rostek (2017) to provide a characterization of the unique robust linear Nash equilibrium in the D2D double auction game. Given his asset holdings χ_i, agent i's objective is to choose the trade size Q_i in the D2D market that maximizes his quadratic utility (6) by choosing the optimal trade size Q_i. The key insight for understanding the nature of strategic trading comes from the observation that the equilibrium demand schedule $Q_i = Q_i(\mathcal{P}^{D2D})$ of dealer i equalizes his marginal utility with his marginal payment for each price,

$$d - \Gamma_i \left(\chi_i + Q_i\right) = \mathcal{P}^{D2D} + \beta_i Q_i,$$

where β_i measures the *price impact* of trader i (also known as the 'Kyle's lambda'; see Kyle (1985)). β_i is the derivative of the inverse residual supply of trader i, which is defined by aggregation through market clearing of the schedules submitted by other traders, $\{Q_j(\mathcal{P}^{D2D})\}_{j\neq i}$. The inverse of price impact is a measure of market liquidity: the lower the price impact, the smaller the price concession a trader needs to accept to trade, the more liquid the market. Importantly, it follows from (4) that *if* trader i knew his price impact

β_i, which is endogenous, he could determine his demand by equalizing his marginal utility and marginal payment pointwise. Let $Q_i(\cdot, \beta_i)$ be the demand schedule defined by (4) for all prices $\mathcal{P}^{D2D}$ by trader i, given his assumed price impact β_i,

$$Q_i(\mathcal{P}^{D2D}, \beta_i) = (\Gamma_i + \beta_i)^{-1}(d - \mathcal{P}^{D2D} - \Gamma_i \chi_i^0). \tag{7}$$

In order to pin down equilibrium price impacts, we note that the market clearing condition requires that the price impact assumed by dealer i be equal to the actual slope of his inverse residual supply, resulting from the aggregation of the other traders' submitted schedules. Proposition 2 (Proposition 1 in Malamud and Rostek (2017)) shows that the system for equilibrium price impacts can be solved explicitly.[11]

Proposition 2 (Centralized Market Equilibrium) *A profile of demand schedules and price impacts $\{Q_i(\cdot, \beta_i), \beta_i\}_i$ is an equilibrium if and only if*

(i) each trader i submits schedule (7), given his price impact β_i,

(ii) trader i's price impact is

$$\beta_i = \left(\sum\nolimits_{j \neq i} (\Gamma_j + \beta_j)^{-1}\right)^{-1}, \ i = 1, \cdots, I.$$

Furthermore,

(iii) Equilibrium exists and is unique.

(iv) Trader i's price impact β_i is given by

$$\beta_i = \frac{2\Gamma_i}{\Gamma_i \mathcal{B} - 2 + \sqrt{(\Gamma_i \mathcal{B})^2 + 4}},$$

[11] For symmetric risk aversions, the equilibrium of Proposition 2 coincides with the equilibrium in Rostek and Weretka (2011), which in turn coincides with Kyle (1989), without nonstrategic traders and assuming independent values). The case of symmetric risk aversions has also been studied in Vayanos (1999) and Vives (2011).

where $\mathcal{B} \in \mathbb{R}_+$ is the unique positive solution to

$$\sum_j (\Gamma_j \mathcal{B} \ + 2 + \sqrt{(\Gamma_j \mathcal{B})^2 + 4})^{-1} = 1/2\,.$$

- β_i *is monotone decreasing in* Γ_i

Corollary 3 characterizes equilibrium trades and prices in the D2D market as a function of LP risk aversions Γ_i and LP post-D2C inventories χ_i (see equation (2)).

Corollary 3 *The D2D market clearing price is given by*

$$\mathcal{P}^{D2D} = d - \mathbf{Q}^* \ \textit{with} \ \mathbf{Q}^* \ \equiv \ \mathcal{B}^{-1} \sum_{j=1}^{M} (\Gamma_j + \beta_j)^{-1} \Gamma_j \chi_j\,. \tag{8}$$

The equilibrium post- D2D trade inventory[12] *of dealer i is*

$$\tilde{\chi}_i \ = \ (\Gamma_i + \beta_i)^{-1} \mathbf{Q}^* + (\Gamma_i + \beta_i)^{-1} \beta_i \chi_i\,. \tag{9}$$

Dealers' equilibrium indirect utility (6) is given by

$$\mathcal{U}_i \ = \ E\left[\chi_i d - 0.5\Gamma_i \chi_i^2 \ + \ (0.5\Gamma_i + \beta_i)(\Gamma_i + \beta_i)^{-2} (\mathbf{Q}^* - \Gamma_i \chi_i)^2\right]\,. \tag{10}$$

Formula (8) is a version of CAPM in the D2D market, but with the *aggregate risk portfolio* $\mathbf{Q}^*$ replacing the standard, equal weighted market portfolio. Each trader's allocation (9) is a combination of dealers' initial inventory χ_i and the aggregate risk portfolio $\mathbf{Q}^*$, which is common to all traders. If the D2D market were competitive, the common portfolio would be the only risk agents were exposed to in equilibrium. Like in the competitive market, $\mathbf{Q}^*$ represents risk that is not diversified in equilibrium and creates a risk premium in prices, $\mathbf{Q}^* = d - \mathcal{P}^{D2D}$. With noncompetitive traders, however, the aggregate risk portfolio is a

[12] See (3).

function of equilibrium price impact. This is intuitive: what constitutes nondiversifiable risk in the market depends on market liquidity.

The central consequence of Corollary 3 is that when agents' risk preferences are heterogeneous, *the aggregate risk* $\mathbf{Q}^*$, which is endogenous, differs from the market portfolio, $\sum_i \chi_i$. In particular, the aggregate risk depends on the joint distribution of endowment risk $\{\chi_i\}_i$ and equilibrium price impact $\{\beta_i\}_i$ rather than on the aggregate inventory alone.

The second main observation is that the D2D market allocates risk in a particular way. Namely, by Proposition 2, *less risk averse agents have greater price impact*: if $\Gamma_1 < ... < \Gamma_I$, then $\beta_1 > \cdots > \beta_I$; less risk averse agents face a more risk averse residual market, and therefore a less elastic residual supply.

5 Customers' Problem

In this section, we study customers' optimization problem. Given the M different price schedules offered by the dealers, a customer optimally decides on how much to trade with each dealer. Substituting dealers' price schedules into customers utility (4), we get that customers' optimization problem takes the form

$$\max_{\{q_{j,i}\}_{i=1}^M} \left\{ -\sum_{i=1}^M q_{j,i} p_i(q_{j,i}) \ + \ \left(\sum_{i=1}^M q_{j,i} + \theta_j \right) d - 0.5\gamma \left(\sum_{i=1}^M q_{j,i} + \theta_j \right)^2 \right\} ,$$

taking dealers' schedules $p_i(q_{j,i}) = \alpha_i + \lambda_i^{-1} q_{j,i}$ as given. Define $a_i = \lambda_i(d - \alpha_i)$ to be the bid-ask spread-normalized deviation from the fundamental value d of the mid-price quoted by dealer i. Let also $a_{-i} = \sum_{j \neq i} a_j$ be the total liquidity-adjusted mid-quote deviation of dealers $j \neq i$, and let $\lambda_{-i} = \sum_{j \neq i} \lambda_j$ be the total liquidity provided by these dealers.

Writing down the first order conditions, we arrive at the following result.

Lemma 4 *The optimal demand of customer j is given by $q_j = \{q_{j,i}\}_{i=1}^M$ with*

$$q_{j,i} \; = \; \delta_i(a_i, \lambda_i, \lambda_{-i}, a_{-i}) \; + \; \eta_i(a_i, \lambda_i, \lambda_{-i}, a_{-i})\, \theta_j \tag{11}$$

with

$$\begin{aligned} \delta_i \; &= \; \frac{0.5a_i + 0.25\gamma(a_i\lambda_{-i} - a_{-i}\lambda_i)}{1 + 0.5\gamma(\lambda_i + \lambda_{-i})} \\ \eta_i \; &= \; -\frac{0.5\gamma\lambda_i}{1 + 0.5\gamma(\lambda_i + \lambda_{-i})} \;, \; i = 1, \cdots, M\,. \end{aligned}$$

The intuition behind the optimal order splitting strategy of Lemma 4 is as follows: Ideally, the customer would like to buy from the dealer with the lowest mid-quote and the highest offered liquidity. Thus, his average demand addressed to dealer i is increasing in a_i and is decreasing in a_j (the attractiveness of trading with other dealers). Customer's demand curve is naturally downward sloping in his inventory (that is, $\eta_i < 0$ for all i), and the size of the slope is proportional to the liquidity λ_i offered by dealer i, as well as to customers' cost of holding inventory, γ.

6 Dealers' Optimal Price Schedules and Equilibrium in the D2C Market

At time $t = 1$, dealer i selects the optimal price schedule $p_i(q)$ that he quotes to all customers, taking as given other dealers' price schedules as well as customers' optimal response (Lemma 4). Substituting (10) into (5), we get that dealers' objective it to maximize

$$\begin{aligned} U_i(a_i, \lambda_i; a_{-i}, \lambda_{-i}) \; &= \; E\Bigg[\sum_{j=1}^{n} p_i(q_{i,j})q_{j,i} \; + \; \chi_i d - 0.5\Gamma_i\chi_i^2 \\ &+ \; (0.5\Gamma_i + \beta_i)(\Gamma_i + \beta_i)^{-2} \left(\mathcal{B}^{-1} \sum_{j=1}^{M} (\Gamma_j + \beta_j)^{-1}\Gamma_j\chi_j - \Gamma_i\chi_i \right)^2 \Bigg] \end{aligned} \tag{12}$$

over a_i, λ_i subject to (2) and (11). That is, dealer's objective is to maximize the total expected price $\Pi_i(Q_i)$ that he receives from his clients, plus the expected autarky utility $\chi_i d - 0.5\Gamma_i \chi_i^2$ from holding the inventory χ_i, plus the surplus from trade in the inter-dealer market. Importantly, dealers' choice of the price schedule characteristics a_i, λ_i influence both his direct revenues, the autarky utility, *and also the gains from trade in the D2D market.* The latter effect is particularly subtle, and also consists of two components: First, it is the dealer's impact on equilibrium price; second, it is the impact of dealer's liquidity provision in the D2C market on the inventories of all other dealers: Indeed, the more liquidity the dealer provides in the D2C market, the less will clients trade with other dealers', directly influencing other dealers' inventories $\chi_j,\ j \neq i$, and hence also influencing the surplus from trade in (13).

Let

$$\Psi \ = \ \left(\frac{\Gamma_1}{\mathcal{B}(\Gamma_1+\beta_1)}, \cdots, \frac{\Gamma_M}{\mathcal{B}(\Gamma_M+\beta_M)}\right)^T$$

be the vector of weights that define the "aggregate risk" in the D2D market, and let also $\Psi_i \ \equiv \ \Psi - \Gamma_i \mathbf{1}_{j=i}$. Denote also $\boldsymbol{\delta} \equiv (\delta_i)_{i=1}^M,\ \boldsymbol{\eta} \equiv (\eta_i)_{i=1}^M$ to be the vectors of coefficients of customers' demand (see (11)). Evaluating the expectation in (13), we get the following expression for dealers' indirect utility.

Lemma 5 *We have*

$$\begin{aligned} U_i(\alpha_i, \lambda_i; \alpha_{-i}, \lambda_{-i}) \ &= \ x_i d - n\lambda_i^{-1} a_i(\delta_i + \eta_i \bar{\theta}) + n\lambda_i^{-1}\Big[(\delta_i + \eta_i\bar{\theta})^2 + \eta_i^2 \sigma_\theta^2\Big] \\ &- 0.5\Gamma_i\Big[(x_i - n\delta_i - n\bar{\theta}\eta_i)^2 + n\sigma_\theta^2 \eta_i^2\Big] \\ &+ \ (0.5\Gamma_i + \beta_i)(\Gamma_i + \beta_i)^{-2}\Big[\Big(\Psi_i \cdot (x \ - \ n\boldsymbol{\delta} - n\bar{\theta}\boldsymbol{\eta})\Big)^2 + n\sigma_\theta^2 (\Psi_i \cdot \boldsymbol{\eta})^2\Big] \end{aligned} \tag{13}$$

and a Nash equilibrium is a collection of $(\alpha_i, \lambda_i)_{i=1}^M$ *such that, for all* i,

$$(\alpha_i, \lambda_i) \;=\; \arg\max_{\alpha_i, \lambda_i} U_i(\alpha_i, \lambda_i; \alpha_{-i}, \lambda_{-i})\,. \tag{14}$$

In the Appendix, we write down the system of first conditions for (14). In general, this system cannot be solved explicitly. However, it is possible to derive analytical approximations to its solution when dealer heterogeneity is small. Such an approximation allows us to capture the first order effects of heterogeneity on equilibrium quantities, while preserving analytical tractability.

7 Small Heterogeneity

Everywhere in the sequel we assume that dealers' heterogeneity is small. We start with the case when dealers do not hold any inventory, and risk aversions are homogeneous, $\Gamma_i = \Gamma$. Then, the following is true.

Proposition 6 *If* $x_i = 0$ *and* $\Gamma_i = \Gamma$ *for all* $i = 1, \cdots, M$, *and* $\bar{\theta} = 0$, *then there exists a unique symmetric equilibrium and the inverse of the price function slope is given by*

$$\lambda^*(\Gamma) \;=\; \frac{(M-2)\gamma - 2\Gamma + \sqrt{((M-2)\gamma - 2\Gamma)^2 + 8\gamma\Gamma(M-1)}}{2\gamma\Gamma(M-1)}\,,$$

while $a_i = 0$ *for all* i. $\lambda^*(\Gamma)$ *is decreasing in both* Γ *and* γ, *and is increasing in* M.

The result of Proposition 6 is very intuitive: Liquidity in the D2C market, as captured by the inverse slope λ^*, is determined by two forces: the willingness (and the ability) of LPs to take on risk (that is, their cost of holding inventory, Γ) and the LPs' market power, determined by their number, M. When Γ increases, or when clients are more aggressively trying to get rid of their inventory (that is, γ is large), LPs optimally widen the spread. At the same time, an increase in M creates a competitive pressure on equilibrium spreads, driving D2C market liquidity up. In particular, in the competitive limit, as $M \to \infty$, we have $\lambda^* \to 1/\Gamma$, consistent with the standard, competitive CAPM whereby price sensitivity to inventory shocks equals the reciprocal of the risk aversion.

Having solved for the equilibrium in the limiting case $x_i = 0$ and $\Gamma_i = \Gamma$, we will now use Taylor approximation to compute the equilibrium for the case when inventories (x_i), mean inventory shock $\bar{\theta}$, and heterogeneity in Γ_i are sufficiently small. In order to state our next result, we introduce the liquidity mis-match, defined as the risk-aversion-weighted average of dealer inventories:

$$X_{mismatch} \equiv \sum_{j=1}^{M} (\Gamma_i - \Gamma^*) x_j \,, \tag{15}$$

where

$$\Gamma^* \equiv \frac{1}{M} \sum_{j=1}^{M} \Gamma_j$$

is the mean risk aversion. By construction, $X_{mismatch}$ is the spread between inventories of high risk aversion and low risk aversion dealers and hence it measures the mismatch between the distribution of inventory risk and their willingness to take on this risk (as captured by Γ_i). While in a perfect market this joint distribution is irrelevant, illiquidity and price-impact imply that the mis-match between risk and risk aversions leads a mis-match between inventory risk and liquidity. This liquidity mis-match matters in both D2C and D2D market segments. The former matters because high risk aversion dealers are less willing to provide liquidity in the D2C market, limiting their ability to efficiently share risk with customers; the latter matters because high risk aversion dealers are less willing to provide liquidity in the D2D market, limiting their ability to efficiently share risk with other dealers. LPs rationally incorporate this mis-match into their D2C quotes. The following is true.

Proposition 7 *Suppose that x_i are sufficiently small and Γ_i are sufficiently close to Γ^*, and* *that $\bar{\theta}$ is sufficiently small. Then, there exists a unique equilibrium with*[13]

$$\lambda_i \;=\; \lambda^*(\Gamma^*) \;+\; \Phi^\Gamma(\Gamma^*)(\Gamma_i - \Gamma^*) \;+\; O(\|\Gamma - \Gamma^*\|^2 + \|x\|^2 + \bar{\theta}^2)\,,$$

where

$$\Phi^\Gamma(x) \;\equiv\; -\frac{x^2(2+\gamma x)(1+(M-1)\gamma x)}{(M^2-2)\,(\gamma x)^2 + 4(M-1)\gamma x + 4} \;<\; 0$$

is negative and increasing in M for $x > 0$ and $M > 1$. Furthermore,

$$\begin{aligned} a_i = \;& \Phi_0^x\, x_i \;+\; \Phi_0^X\, X \;+\; \Phi_0^\Gamma\, X_{mismatch} \;+\; \Phi_0^\theta\, \bar{\theta} \;+\; (\Gamma_i - \Gamma^*)[\Phi_1^x\, x_i + \Phi_1^X\, X + \Phi_1^\theta\, \bar{\theta}] \\ &+\; O((\|\Gamma - \Gamma^*\|^2 + \|x\|^2 + \bar{\theta}^2)^{3/2})\,, \end{aligned} \tag{16}$$

where

$$X \;\equiv\; \sum_{j=1}^{M} x_j\,,$$

and $\Phi_0^x, \Phi_0^X, \Phi_0^\Gamma, \Phi_0^\theta, \Phi_1^x, \Phi_1^X, \Phi_1^\Gamma, \Phi_1^\theta$ are rational functions[14] *of $n, M, \gamma,$ and Γ^*. Moreover,*

$$\Phi_0^x \;>\; 0\,.$$

Proposition 7 characterizes the optimal strategic behavior of dealers in the D2C market, given their inventories $\{x_j\}_{j=1}^M$ and holding costs $\{\Gamma_j\}_{j=1}^M$. The fact that Φ^Γ is negative means that a high holding cost (risk aversion) dealer is less willing to provide liquidity and hence sets a wider spread λ_i^{-1} in the D2C market. The fact that Φ_0^x is positive means that dealers

[13] As usual, we use $\|x\| = \left(\sum_i x_i^2\right)^{1/2}$ to denote the Euclidean norm.

[14] We provide explicit expressions for these coefficients in the Appendix.

"shade prices" in the D2C market. That is, they shift prices in the direction opposite to their inventory.[15]

The most important consequence of Proposition 7 is that forces of competition make each dealer's quote depend on other dealers' inventories and risk aversions. The term $\Phi_0^X\, X$ shows that the average inventory matters. The term $\Phi_0^\Gamma\, X_{mismatch}$ shows how heterogeneity in risk aversions impacts quotes. This is the first time we see the liquidity mis-match, $X_{mismatch}$, appear in the expressions. This term originates from dealers' expectations about the D2D trade, characterized in Corollary 3.

8 The Identification Problem

Our goal is to use Proposition 7 in order to identify the actual (directly unobservable) dealers' characteristics $\{\Gamma_i,\ x_i\}_{i=1}^M$. First, inverting the expression in Proposition 6, we get that, absent heterogeneity, a bid-ask spread λ^{-1} in the D2C market implies that dealers' have a holding cost of

$$\Gamma^*(\lambda)\ \equiv\ \frac{1}{\lambda}\,\frac{2+(M-2)\gamma\lambda}{2+(M-1)\gamma\lambda}\,. \tag{17}$$

Now, given (17), we can recover dealers' characteristics from the observed prices and spreads when heterogeneity is small. The following result is a direct consequence of Proposition 7.

Proposition 8 *Suppose that a_i are sufficiently small and the dispersion in bid-ask spreads λ_i^{-1} is sufficiently small, and that $\bar\theta$ is sufficiently small. Then,*

$$\Gamma_i\ =\ \Gamma^*(\lambda^*)\ +\ \varphi^\lambda(\lambda^*)(\lambda_i-\lambda^*)\ +\ O(\|\lambda-\lambda^*\|^2+\|a\|^2+\bar\theta^2) \tag{18}$$

[15] Recall that a_i is negatively related to the mid-quote: $a_i = \lambda_i(d-\alpha_i)$.

where

$$\lambda^* \;=\; \frac{1}{M}\sum_{i=1}^{M}\lambda_i \qquad \text{and} \qquad \varphi^{\lambda}(\cdot) \;=\; \frac{1}{\Phi^{\Gamma}(\cdot)}\,.$$

Furthermore,

$$\begin{aligned} x_i = \;& \varphi_0^a a_i \;+\; \varphi_0^A A \;+\; \varphi_{0,\lambda}^A A_{mismatch} \;+\; \varphi_0^{\theta}\bar{\theta} \;+\; (\lambda_i-\lambda^*)[\varphi_1^a a_i+\varphi_1^A A+\varphi_{1,\lambda}^A A^{\lambda}+\varphi_1^{\theta}\bar{\theta}] \\ & \;+\; O((\|\Gamma-\Gamma^*\|^2+\|x\|^2+\bar{\theta}^2)^{3/2})\,, \end{aligned} \tag{19}$$

where

$$A \;=\; \sum_{j=1}^{M} a_j; \qquad A_{mismatch} \;=\; \sum_{j=1}^{M}(\lambda_j-\lambda^*)a_j,$$

and $\varphi_0^a, \varphi_0^A, \varphi_{0,\lambda}^A, \varphi_0^{\theta}, \varphi_1^a, \varphi_1^A, \varphi_{1,\lambda}^A,$ *and* φ_1^{θ} *are rational functions of* n, M, γ *and* λ^*. *Moreover,*

$$\varphi_0^a(\cdot) \;=\; 1/\Phi_0^x(\cdot) \;>\; 0\,.$$

$$\begin{aligned} x_i = \;& \varphi_0^a a_i \;+\; \varphi_0^A A \;+\; \varphi_{0,\lambda}^A A_{mismatch} \;+\; \varphi_0^{\theta}\bar{\theta} \;+\; (\lambda_i-\lambda^*)[\varphi_1^a a_i+\varphi_1^A A+\varphi_{1,\lambda}^A A^{\lambda}+\varphi_1^{\theta}\bar{\theta}] \\ & \;+\; O((\|\Gamma-\Gamma^*\|^2+\|x\|^2+\bar{\theta}^2)^{3/2})\,, \end{aligned}$$

Proposition 8 is key to our subsequent analysis. It shows how dealers' risk aversions (see (18)) and inventories (see (19)) can be uniquely recovered from the observed quotes and bid-ask spreads. In particular, we see how the liquidity mis-match (15) from Proposition 7 gets mapped in an analogous object in the space of quotes: $A_{mismatch}$, given by the bid-ask spread weighted average of transformed quotes, $a_i = \lambda_i(d-\alpha_i)$. Since, by (16), a_i is positively related to x_i, while λ_i are negatively related to λ_i, we get that $A_{mismatch}$ is negatively related to the liquidity mismatch, $X_{mismatch}$.

In the next section we will substitute these recovered quantities into formula (8) for the equilibrium price in the D2D market in order to get an explicit link between prices and spreads in the two market segments. However, we believe that the results of Proposition 8 may also be of independent interest for regulators who could potentially use these expressions to monitor in real time the dynamics of liquidity in the foreign exchange markets.

9 Dynamics and Empirical Predictions: From D2C to D2D Prices

In this section, we test the basic predictions of our model empirically. To this end, we assume that, over a sufficiently small, period (such as, e.g., a few seconds), no new fundamental information arrives, and hence prices in both the D2C and the D2D markets are driven exclusively by inventory shocks. We assume that all other model parameters including LPs' risk aversions $\Gamma_i,\ i = 1, \cdots, M$ stay constant over that time horizon.[16] Under this assumption, formulas of Proposition 8 allow us to identify the unobservable parameter vectors (Γ_i) and (x_i) from the observable vectors of mid-prices, α_i, and bid-ask spreads, b_i, in the D2C market.

Substituting optimal client trades from (11) into formula (2) for the post-D2C trading inventories of LPs, we can get an expression for those inventories as a function of LPs initial inventories x_i, as well as D2C quotes (α_i, λ_i),

$$\chi_i \ = \ \chi_i(x_i, \alpha_i, \lambda_i, \alpha_{-i}, \lambda_{-i})\,.$$

Now, using formula (19) from Proposition 8, we can re-express initial inventories $x_i = x_i(\alpha_i, \lambda_i, \alpha_{-i}, \lambda_{-i})$ as a function of D2C quotes (α_i, λ_i), which allows us to finally get dealers'

[16] We do not micro-found the origins of the shocks to LPs risk aversions Γ. These shocks are typically coming from banks capital requirements, shocks to LPs cost of fundings, and regulatory constraint that may be binding at a bank level and transmit into individual trading desk behaviour at the bank.

post-D2C inventories as a function of D2C quotes:

$$\chi_i \;=\; \chi_i(x_i(\alpha_i, \lambda_i, \alpha_{-i}, \lambda_{-i}), \alpha_i, \lambda_i, \alpha_{-i}, \lambda_{-i}) \tag{20}$$

See Lemma 13 in the Appendix for an explicit expression. We can now substitute post-D2C inventories (20) into the formula in Corollary 3 and compute the equilibrium price in the D2D market following the D2C trading round. This brings us to the following result.

Proposition 9 *The price $P = P^{D2D}$ in the D2D market is linked to prices in the D2C market via*

$$\begin{aligned}\mathcal{P}^{D2D} \approx\; & \left[1+\frac{2\gamma\lambda^*}{(2+M\gamma\lambda^*)[2+(M-1)\gamma\lambda^*]}\right](\bar{\alpha}+\hat{\alpha}) \\ & +\; (\lambda^*)^2\left[\Gamma^*(M\varphi^A_{0,\lambda}+\varphi^a_1) \;+\; \frac{M\varphi^\lambda(\varphi^a_0-0.5n)}{M^2-2M+2}\right]\hat{\alpha} \\ & -\; \frac{2\gamma\lambda^*}{(2+M\gamma\lambda^*)[2+(M-1)\gamma\lambda^*]}d \;-\; \Gamma^*\varphi^\theta_0\bar{\theta} \;-\; \gamma\frac{1}{2+\gamma\Lambda}\lambda^*\Gamma^*\Theta\,.\end{aligned}$$

Proposition 9 provides an explicit expression linking the price in the D2D market to quotes in the D2C market, as well as client inventory shocks, Θ.[17] The most important consequence of Proposition 9 is that, in the presence of heterogeneity, the joint cross-sectional distribution of quotes and bid-ask spreads matters for D2D trading. Namely, in addition to the average mid-quote $\bar{\alpha}_i$, the price also depends on the bid-ask spread-weighted mid-price

$$\hat{\alpha} \;=\; \frac{1}{\sum_{j=1}^M \lambda_j}\sum_{j=1}^M(\lambda_i-\lambda^*)\alpha_i\,,$$

which is effectively a spread between low bid-ask spread (large λ_i) mid-quotes and high bid-ask spread (low λ_i) quotes. Since prices are negatively related to inventories, $\hat{\alpha}$ is negatively related to the liquidity mis-match.

[17] Note that, while the coefficient on the fundamental value, d, is negative, the price $\mathcal{P}^{D2D}$ depends on positively on d (at it should) due to a positive link between α_i and d.

By Proposition 9, the sensitivity of $\mathcal{P}^{D2D}$ to $\hat{\alpha}$ is given by

$$1 \; + \; \frac{2\gamma\lambda^*}{(2+M\gamma\lambda^*)[2+(M-1)\gamma\lambda^*]} \; + \; (\lambda^*)^2 \left[\Gamma^*(M\varphi^A_{0,\lambda}+\varphi^a_1) \; + \; \frac{M\varphi^\lambda(\varphi^a_0-0.5n)}{M^2-2M+2}\right].$$

The sign and the magnitude of this coefficient is driven may many competing forces in the two market segments. While do not have analytical results about the sign of this coefficient, extensive numerical simulations indicate that $\hat{\alpha}$ has a positive sign. Furthermore, consistent with the standard CAPM story, the coefficient on customer inventory, Θ, is always negative. We formalize these results in the following predictions.

Prediction 1. $\mathcal{P}^{D2D}$ is positively related to $\hat{\alpha}$;

Prediction 2. $\mathcal{P}^{D2D}$ is negatively related to Θ.

We now discuss the link between bid-ask spreads in the D2D and the D2C market segments. The most natural measure of bid-ask spreads in our theoretical double auction model of the D2D market is price impact, $2\beta_i$. Indeed, as in the standard Kyle (1989) model, $2\beta_i$ is the spread between the price at which agent i can sell one unit of the asset and the price at which he can buy one unit of the asset. The following is true.

Proposition 10 *Equilibrium price impacts β_i in the D2D market are monotone decreasing in the dispersion of D2C bid-ask spreads,*

$$\frac{1}{M}\sum_i(\lambda_i-\lambda_*)^2 \,.$$

The result of Proposition 10 is a consequence of a convexity effect: Dealers' liquidity provision (as captured by the slope of their demand schedule in the D2D market) is inversely proportional to their risk aversion, and total D2D market liquidity is given by the total slope of dealers' demand. Since the function $1/x$ is convex, total slope increases in the dispersion of risk aversions, which is in turn proportional to the dispersion in bid-ask spreads. We formalize the result of Proposition 10 in our last prediction.

Prediction 3. Bid-ask spreads in the D2D market are negatively related to the dispersion of D2C bid-ask spreads.

We test Predictions 1-3 in the next section.

10 Empirical analysis

10.1 Empirical specifications

We test the model's key prediction regarding the joint dynamics of the D2D and D2C mid-prices. Proposition 9 predicts the following relation between D2C and D2D prices:

$$\mathcal{P}_{t+1}^{D2D} = a_0 + a_1\bar{\alpha}_t + a_2\hat{\alpha}_t + a_3\Theta_t + a_4\bar{\theta}_t \, ,$$

where

$$a_1, \, a_2 > 0 \, .$$

We test this relation by regressing D2D prices on both $\bar{\alpha}$ and $\hat{\alpha}$, and a set of controls. We replace dependent and independent variables with their first differences in our main empirical specification because prices are know to be persistent. Following the extant literature, we also replace prices by log prices (see for example Evans and Lyons (2002b)). The results are qualitatively similar when we use prices instead. Thus, our main empirical specification is:

$$\Delta p_{t+\ell,\,\ell}^{D2D} = a_0 + a_1\Delta\bar{\alpha}_{t,\,\ell} + a_2\Delta\hat{\alpha}_{t,\,\ell} + controls_t \, ,$$

where

$$X_{t,\ell} = X_t - X_{t-\ell} \qquad \text{and} \qquad p_t^{D2D} = \log \mathcal{P}_t^{D2D} \, ,$$

for several lags ℓ ranging from 1 to 60 seconds. We include four controls in our analysis. The first is the liquidity demand of customers $ProxyLiqShock$, which is a proxy for effects of Θ and $\bar{\theta}$ on prices. The other three are various measures of the realized order flow in the D2D market, since Evans and Lyons (2002b) show that this variable is contemporaneously related to price changes in the D2D market.

10.2 Data description

We test the model's predictions using three high frequency datasets. The first two datasets are provided to us the Swissquote bank, one of the top on-line FOREX brokers in the world. The first dataset contains price schedules for the currency pair "EUR/USD"[18] submitted to Swissquote by the largest FOREX dealers over the period from the 16th of May 2016 to the 1st of June 2016. Price schedules are stepwise functions mapping volumes to bid/ask quotes. The second dataset contains the orders submitted for execution to Swissquote by it's clients for the same currency pair and the same time period. The third dataset is the EBS dataset provided to us by NEX Data. This dataset is a comprehensive account of FX best bid and ask aggregated within each second. All data are time-stamped with a 1 second precision.

For the Swissquote datasets, we filter the original data sample on quotes so as to retain price schedules from a subsample of dealers who were simultaneously providing quotes during a long continuous time interval within each trading day. After applying such a filtering procedure, we retain quotes from ten large dealers, which yields approximatively 864000 observations per trading day. We then estimate the mid-price ($\alpha_{j,t}$) and the marginal half-spread ($b_{j,t}$) for each dealer and each time period. We then constructs our dependent variables using these estimated values.

Our proxy for the customers' liquidity demand is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure). This

[18] According to the Trennial Central Bank Survey (April 2016), this currency pair accounts for the largest turnover on the FOREX market.

definition follows from our assumption that any online broker acts as an intermediary between retail clients and large dealers, and such brokers are competitive. Then, the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure) can be seen as *a scaled proxy* of total customer shocks of $\bar{\Theta}_t$. However, clients' orders arrive at a much lower frequency compared to the frequency at which dealers update their quotes, so that the obtained time-series of the proxy of $\bar{\Theta}$ is very sparse. To mitigate this problem, we smooth the original time series of the aggregate clients' order flow by taking a 5-minute moving average.

Finally, our proxy for the D2D price is the mid-price in the inter-dealer market, which we compute using the EBS dataset. We construct order flow in the D2D market using the EBS dataset, which provides volume and buy/sell information for each executed transaction, aggregated at the second level. We compute the following proxies for the D2D order flow:

- **OF_d2d_Vol:** Aggregate daily (signed) order flow in the D2D market.
- **OF_d2d_Bin:** Aggregate daily binary Buy and Sell executed in the D2D market, where a buy order is order marked as $+1$ and a Sell order as -1. This measure replicates the approach of Evans and Lyons (2002b).
- **OF_Vol_1h:** Aggregate (signed) order flow in the D2D market over the previous 1 hour.
- **OF_Vol_24h:** Aggregate (signed) order flow in the D2D market over the previous 24 hours.

We report the summary statistics for the final merged dataset in Table 1.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	5%	50%	95%
α	1.1245	0.0081	1.1098	1.1416	1.1133	1.1224	1.1371
b	0.00002	0.00003	0	0.00099	0	0.00001	0.00005
$\bar{\Theta}$	37	5667	-210403	138445	-6413	0	6500
$\bar{\alpha}$	1.1245	0.0081	1.1098	1.1416	1.1133	1.1223	1.1371
$\hat{\alpha}$	0.000001	0.000019	-0.001952	0.000732	-0.000006	0	0.000008

10.3 Empirical results

We start with a test of the first two predictions of our model, that is that our novel measure of systematic risk positively predicts future FX exchange rates while customers' order flow negatively predicts those rates. We regress changes in D2D log prices on both lagged changes in D2C prices and lagged clients' order flow proxies, with the lag $\ell = 10$ seconds. Table 2 reports the results of our main the estimation for the ten seconds lag. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis. The first column shows that our measure of cross-sectional liquidity-risk mismatch in the D2C market, $\hat{\alpha}$, positively and significantly forecasts price fluctuations in the D2D market. In terms of economic significance, an increases in the cross-sectional liquidity-risk mismatch in the D2C market by 1% results in a 2.9% increase in the FX exchange rate. Column 2 shows a striking result: Customers' order flow in the D2C market negatively forecasts price changes in the D2D market. The negative sign means that an increased in the customers' liquidity demand results in future lower price. However, the economic significance is less than that of our novel measure of systematic risk. The remaining columns show that controlling for lagged D2D order flow does not change our results.

We repeat this empirical test for lags $\ell = 1, 5, 15, 20, 30, 40, 45, 50, 60$ and present the results in Tables 4 to 13. There are two main observations from the tables. First, our measure of systematic risk in the D2D market is significant for small lags ℓ and looses significance as we increase ℓ. This result is intuitive: The heterogeneity among dealers, which determines our measure of systematic risk, is a function of dealers' liquidity shocks, shocks that are not persistent. Second, the predictive power of the D2C customers' liquidity demand ($ProxyLiqShock$) is significant over longer horizons relative to both our measure of systematic risk and the average mid-price in the D2C market.

[**Include Table 2 here.**]

References

Bacchetta, Philippe and Eric Van Wincoop, "Can information heterogeneity explain the exchange rate determination puzzle?," *The American Economic Review*, 2006, *96* (3), 552–576.

Back, Kerry and Shmuel Baruch, "Strategic liquidity provision in limit order markets," *Econometrica*, 2013, *81* (1), 363–392.

Banti, Chiara, Kate Phylaktis, and Lucio Sarno, "Global liquidity risk in the foreign exchange market," *Journal of International Money and Finance*, 2012, *31* (2), 267–291.

Berger, David W, Alain P Chaboud, Sergey V Chernenko, Edward Howorka, and Jonathan H Wright, "Order flow and exchange rate dynamics in electronic brokerage system data," *Journal of international Economics*, 2008, *75* (1), 93–109.

Biais, Bruno, David Martimort, and Jean-Charles Rochet, "Competing mechanisms in a common value environment," *Econometrica*, 2000, *68* (4), 799–837.

___ , ___ , and ___ , "Competing mechanisms in a common value environment: A Corrigendum," *Econometrica*, 2013, *81* (4), 393–406.

Bjønnes, Geir H, Carol L Osler, Dagfinn Rime, and Norges Bank, "Sources of information advantage in the foreign exchange market," *Norges Bank*, 2011, p. 46.

Bjønnes, Geir Høidal and Dagfinn Rime, "Dealer behavior and trading systems in foreign exchange markets," *Journal of Financial Economics*, 2005, *75* (3), 571 – 605.

___ , ___ , and Haakon O Aa Solheim, "Liquidity provision in the overnight foreign exchange market," *Journal of International Money and Finance*, 2005, *24* (2), 175–196.

Brandt, Michael W and Kenneth A Kavajecz, "Price discovery in the US Treasury market: The impact of orderflow and liquidity on the yield curve," *The Journal of Finance*, 2004, *59* (6), 2623–2654.

Breedon, Francis and Paolo Vitale, "An empirical study of portfolio-balance and information effects of order flow on exchange rates," *Journal of International Money and Finance*, 2010, *29* (3), 504–524.

Brunnermeier, Markus K, Stefan Nagel, and Lasse H Pedersen, "Carry trades and currency crashes," *NBER macroeconomics annual*, 2008, *23* (1), 313–348.

Cao, H Henry, Martin D Evans, Richard K Lyons et al., "Inventory Information," *The Journal of Business*, 2006, *79* (1), 325–364.

Collin-Dufresne, Pierre, Benjamin Junge, and Anders B Trolle, "Market Structure and Transaction Costs of Index CDSs," 2016.

Danielsson, Jon and Ryan Love, "Feedback trading," *International Journal of Finance & Economics*, 2006, *11* (1), 35–53.

___ , Jinhui Luo, and Richard Payne, "Exchange rate determination and inter-market order flow effects," *The European Journal of Finance*, 2012, *18* (9), 823–840.

Dunne, Peter G, Harald Hau, and Michael Moore, "A tale of two platforms: dealer intermediation in the European sovereign bond market," 2010.

Evans, Martin DD, "FX trading and exchange rate dynamics," *The Journal of Finance*, 2002, *57* (6), 2405–2447.

___ and Dagfinn Rime, "Order flow information and spot rate dynamics," *Journal of International Money and Finance*, 2016, *69*, 45–68.

___ and Richard K Lyons, "Informational integration and FX trading," *Journal of International Money and Finance*, 2002, *21* (6), 807–831.

___ and ___ , "Order flow and exchange rate dynamics," *Journal of political economy*, 2002, *110* (1), 170–180.

___ and ___ , "Do currency markets absorb news quickly?," *Journal of International Money and Finance*, 2005, *24* (2), 197–217.

___ and ___ , "Meese-Rogoff Redux: Micro-Based Exchange-Rate Forecasting," *American Economic Review*, 2005, pp. 405–414.

Evans, Martin D.D. and Richard K. Lyons, "A New Micro Model of Exchange Rate Dynamics," Working paper, Georgetown University 2005.

Evans, Martin DD et al., *Exchange-Rate Dynamics*, Princeton University Press, 2011.

Fan, Mintao and Richard K Lyons, "Customer trades and extreme events in foreign exchange," *Monetary History, Exchange Rates and Financial Markets: Essays in Honour of Charles Goodhart*, 2003, pp. 160–179.

Frankel, Jeffrey A, Giampaolo Galli, and Alberto Giovannini, *The microstructure of foreign exchange markets*, University of Chicago Press, 2009.

Froot, Kenneth A and Tarun Ramadorai, "Currency Returns, Intrinsic Value, and Institutional-Investor Flows," *The Journal of Finance*, 2005, *60* (3), 1535–1566.

Gabaix, Xavier and Matteo Maggiori, "International liquidity and exchange rate dynamics," *The Quarterly Journal of Economics*, 2015, *130* (3), 1369–1420.

Gargano, Antonio, Steven J. Riddiough, and Lucio Sarno, "The Value of Volume in Foreign Exchange," Working Paper, University of Melbourne 2018.

Hau, Harald, William Killeen, and Michael Moore, "The euro as an international currency: explaining puzzling first evidence from the foreign exchange markets," *Journal of International Money and Finance*, 2002, *21* (3), 351–383.

Ho, Thomas and Hans R Stoll, "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial economics*, 1981, *9* (1), 47–73.

Killeen, William P, Richard K Lyons, and Michael J Moore, "Fixed versus flexible: Lessons from EMS order flow," *Journal of International Money and Finance*, 2006, *25* (4), 551–579.

King, Michael, Lucio Sarno, and Elvira Sojli, "Timing exchange rates using order flow: The case of the Loonie," *Journal of Banking & Finance*, 2010, *34* (12), 2917–2928.

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

___ , "Informed speculation with imperfect competition," *The Review of Economic Studies*, 1989, *56* (3), 317–355.

Lyons, Richard K, "Tests of microstructural hypotheses in the foreign exchange market," *Journal of Financial Economics*, 1995, *39* (2), 321–351.

___ **et al.**, *The Microstructure Approach to Exchange Rates*, Vol. 1, The MIT Press, 2001.

Madhavan, Ananth and Seymour Smidt, "An analysis of changes in specialist inventories and quotations," *The Journal of Finance*, 1993, *48* (5), 1595–1628.

Malamud, Semyon and Andreas Schrimpf, "An Intermediation-Based Model of Exchange Rates," Working Paper, BIS 2017.

___ **and Marzena Rostek**, "Decentralizaed Exchange," *American Economic Review*, 2017, *107* (11), 3320–3362.

Mancini, Loriano, Angelo Ranaldo, and Jan Wrampelmeyer, "Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums," *The Journal of Finance*, 2013, *68* (5), 1805–1841.

Meese, Richard A and Kenneth Rogoff, "Empirical exchange rate models of the seventies: Do they fit out of sample?," *Journal of international economics*, 1983, *14* (1-2), 3–24.

Michaelides, Alexander, Andreas Milidonis, and George Nishiotis, "Private Information in Currency Markets," Technical Report, Journal of Financial Economics (forthcoiming) 2018.

Moore, Michael, Andreas Schrimpf, and Vladyslav Sushko, "Downsized FX markets: causes and implications," *BIS Quaterly Review*, December 2016.

Osler, Carol L, Alexander Mende, and Lukas Menkhoff, "Price discovery in currency markets," *Journal of International Money and Finance*, 2011, *30* (8), 1696–1718.

Rime, Dagfinn, Lucio Sarno, and Elvira Sojli, "Exchange rate forecasting, order flow and macroeconomic information," *Journal of International Economics*, 2010, *80* (1), 72–88.

Rostek, Marzena and Marek Weretka, "Dynamic thin markets," *The Review of Financial Studies*, 2015, *28* (10), 2946–2992.

Vayanos, Dimitri, "Strategic trading and welfare in a dynamic market," *The Review of Economic Studies*, 1999, *66* (2), 219–254.

Vives, Xavier, "Strategic supply function competition with private information," *Econometrica*, 2011, *79* (6), 1919–1966.

A Proofs

Proof of Lemma 4. The first order condition is

$$-\alpha_i - 2\lambda_i^{-1} q_i + d - \gamma(\sum_i q_i + \theta) \;=\; 0$$

where $q = (q_i)$ and this gives

$$q \;=\; (2B + \gamma \mathbf{1})^{-1}((d - \gamma\theta)\mathbf{1} - \alpha)$$

where $B = \mathrm{diag}(\lambda_i^{-1}),\ \alpha = (\alpha_i)$.

Q.E.D.

Proof of Lemma 5. We have

$$E[\Pi_i(Q_i)] \;=\; E[\sum_j ((d-\lambda_i^{-1}a_i)+\lambda_i^{-1}(\delta_i+\eta_i\theta_j))(\delta_i+\eta_i\theta_j)] \;=\; (d-\lambda_i^{-1}a_i)n(\delta_i+\eta_i\bar{\theta})+\lambda_i^{-1}(n(\delta_i+\eta_i\bar{\theta})^2+n\eta_i^2\sigma_\theta^2)\,.$$

Thus, the utility becomes

$$\begin{aligned}
E[U_i(\chi_i;(\chi_j)_{j\neq i})] = &\ (d-\lambda_i^{-1}a_i)n(\delta_i+\eta_i\bar{\theta})+\lambda_i^{-1}(n(\delta_i+\eta_i\bar{\theta})^2+n\eta_i^2\sigma_\theta^2)\\
&+(x_i-n\delta_i-\eta_i n\bar{\theta})d-0.5\Gamma_i[(x_i-n\delta_i-\eta_i n\bar{\theta})^2+\eta_i^2 n\sigma_\theta^2]\\
&+\ (0.5\Gamma_i+\beta_i)(\Gamma_i+\beta_i)^{-2}\Big[\Big(\Psi_i\cdot(x\ -\ n\boldsymbol{\delta}-n\bar{\theta}\boldsymbol{\eta})\Big)^2+n\sigma_\theta^2(\Psi_i\cdot\boldsymbol{\eta})^2\Big]\\
=&\ x_i d-n\lambda_i^{-1}a_i(\delta_i+\eta_i\bar{\theta})+n\lambda_i^{-1}\Big[(\delta_i+\eta_i\bar{\theta})^2+\eta_i^2\sigma_\theta^2\Big]\\
&-0.5\Gamma_i\Big[(x_i-n\delta_i-n\bar{\theta}\eta_i)^2+n\sigma_\theta^2\eta_i^2\Big]\\
&+\ (0.5\Gamma_i+\beta_i)(\Gamma_i+\beta_i)^{-2}\Big[\Big(\Psi_i\cdot(x\ -\ n\boldsymbol{\delta}-n\bar{\theta}\boldsymbol{\eta})\Big)^2+n\sigma_\theta^2(\Psi_i\cdot\boldsymbol{\eta})^2\Big].
\end{aligned}$$

where we have defined

$$\Psi\ =\ \left(\frac{\Gamma_1}{\mathcal{B}(\Gamma_1+\beta_1)},\cdots,\frac{\Gamma_M}{\mathcal{B}(\Gamma_M+\beta_M)}\right)^T \qquad \text{and} \qquad \Psi_i\ \equiv\ \Psi-\Gamma_i\mathbf{1}_{j=i}\,.$$

Now, we have

$$\begin{aligned}
\boldsymbol{\delta}\ &=\ \frac{[2+\gamma\Lambda]\mathbf{a}\ -\ \gamma A\boldsymbol{\lambda}}{4+2\gamma\Lambda}\\
\boldsymbol{\eta}\ &=\ -\frac{\gamma\boldsymbol{\lambda}}{2+\gamma\Lambda}\,.
\end{aligned}$$

Q.E.D.

Lemma 11 (FOC)

Proof. It follows from Lemma 4 that

$$\begin{aligned}
\partial_{a_i}\delta_j &= \frac{\mathbb{1}_i(j)[2+\gamma\Lambda] \;-\; \gamma\lambda_j}{4+2\gamma\Lambda} \\
\partial_{a_i}\eta_j &= 0 \\
\partial_{\lambda_i}\delta_j &= -\frac{2\gamma\Big(2a_j+\gamma[a_j(\Lambda-\lambda_j)-(A-a_j)\lambda_j]\Big)}{(4+2\gamma\Lambda)^2} \;+\; \gamma\frac{a_j-A\mathbb{1}_i(j)}{4+2\gamma\Lambda} \\
\partial_{\lambda_i}\eta_j &= \frac{\gamma^2\lambda_j}{(2+\gamma\Lambda)^2} \;-\; \mathbb{1}_i(j)\frac{\gamma}{2+\gamma\Lambda}\,,
\end{aligned}$$

where $\mathbb{1}_i(\cdot)$ is the indicator function. We make the following definition:

$$\begin{aligned}
&\partial_{a_i}\boldsymbol{\delta} \equiv\; (\partial_{a_i}\delta_1,\cdots,\partial_{a_i}\delta_n)^T \;=\; \frac{1}{4+2\gamma\Lambda}\Big[(2+\gamma\Lambda)\boldsymbol{e}_i-\gamma\boldsymbol{\lambda}\Big] \\
&\partial_{a_i}\boldsymbol{\eta} =\; 0 \\
&\partial_{\lambda_i}\boldsymbol{\delta} \equiv\; (\partial_{\lambda_i}\delta_1,\cdots,\partial_{\lambda_i}\delta_n)^T \;=\; -\frac{2\gamma}{(4+2\gamma\Lambda)^2}\Big((2+\gamma\Lambda)\boldsymbol{a}-\gamma A\boldsymbol{\lambda}\Big) \;+\; \frac{\gamma}{4+2\gamma\Lambda}\Big(\boldsymbol{a}-A\boldsymbol{e}_i\Big) \\
&\partial_{\lambda_i}\boldsymbol{\eta} \equiv\; (\partial_{\lambda_i}\eta_1,\cdots,\partial_{\lambda_i}\eta_n)^T \;=\; \frac{\gamma^2}{(2+\gamma\Lambda)^2}\boldsymbol{\lambda} \;-\; \frac{\gamma}{2+\gamma\Lambda}\boldsymbol{e}_i\,,
\end{aligned}$$

where $\boldsymbol{e}_i$ is the ith coordinate vector. Then, the FOCs are

$$\begin{aligned}
0 =\;& n(\partial_{a_i}\delta_i)\Big[\Gamma_i x_i-\lambda_i^{-1}a_i\Big]+n\lambda_i^{-1}\Big[(2-n\Gamma_i\lambda_i)(\partial_{a_i}\delta_i)-1\Big](\delta_i+\eta_i\bar{\theta}) \\
&-\; n(\Gamma_i+2\beta_i)(\Gamma_i+\beta_i)^{-2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\Big(\Psi_i\cdot(x \;-\; n\boldsymbol{\delta}-n\bar{\theta}\boldsymbol{\eta})\Big) \\
0 =\;& n\lambda_i^{-2}a_i(\delta_i+\eta_i\bar{\theta})-n\lambda_i^{-2}\Big[(\delta_i+\eta_i\bar{\theta})^2+\eta_i^2\sigma_\theta^2\Big]+n\lambda_i^{-1}\sigma_\theta^2\Big[2-\Gamma_i\lambda_i\Big](\partial_{\lambda_i}\eta_i)\eta_i \\
&+\lambda_i^{-1}n\Big[2(\delta_i+\eta_i\bar{\theta})+\lambda_i\Gamma_i(x_i-n\delta_i-n\bar{\theta}\eta_i)-a_i\Big](\partial_{\lambda_i}\delta_i+\bar{\theta}\partial_{\lambda_i}\eta_i) \\
&-\; n(\Gamma_i+2\beta_i)(\Gamma_i+\beta_i)^{-2}\Big[\Big(\Psi_i\cdot(\partial_{\lambda_i}\boldsymbol{\delta}+\bar{\theta}\partial_{\lambda_i}\boldsymbol{\eta})\Big)\Big(\Psi_i\cdot(x \;-\; n\boldsymbol{\delta}-n\bar{\theta}\boldsymbol{\eta})\Big)-\sigma_\theta^2(\Psi_i\cdot\partial_{\lambda_i}\boldsymbol{\eta})(\Psi_i\cdot\boldsymbol{\eta})\Big]\,.
\end{aligned}$$

Q.E.D.

Proof of Proposition 6. We present the proof of the monotonicity results.

$$
\begin{aligned}
\frac{\partial \lambda^*}{\partial \gamma} &= \frac{\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M} \; - \; (\gamma M+2\Gamma)}{\gamma^2(M-1)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M}} \\
&= \frac{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M \; - \; (\gamma M+2\Gamma)^2}{\gamma^2(M-1)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M}[\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M} \; + \; (\gamma M+2\Gamma)]} \\
&= \frac{-4\gamma^2(M-1)}{\gamma^2(M-1)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M}[\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M} \; + \; (\gamma M+2\Gamma)]} \\
&< 0
\end{aligned}
$$

since $M \geq 2$.

$$
\frac{\partial \lambda^*}{\partial \Gamma} = \frac{-(M-2)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M} - \gamma(M-2)^2 - 2M\Gamma}{2\Gamma^2(M-1)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M}} \; < \; 0\,.
$$

$$
\frac{\partial \lambda^*}{\partial M} = \frac{\gamma^2(M-2)+(\gamma+2\Gamma)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M} - 2\Gamma(\gamma+2\Gamma+\gamma M)}{2\gamma\Gamma(M-1)^2\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M}}\,.
$$

To show that this derivative is always positive for $M \geq 2$ it is enough to show that its numerator is positive for $M \geq 2$. We do so by showing that the numerator is increasing and positive for $M \geq 2$. Consider the numerator. Its derivative with respect to M is

$$
\gamma\left(\gamma - 2\Gamma + (\gamma+2\Gamma)\frac{2\Gamma+\gamma(M-2))}{\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M}}\right)\,.
$$

We now show that the term in brackets is positive. First,

$$
1 \geq \frac{\gamma(M-2)+2\Gamma}{\sqrt{\gamma^2(M-2)^2+4\gamma\Gamma M+4\Gamma^2}} \; \geq \; 1 - \frac{2\gamma}{\gamma M+2\Gamma}\,.
$$

Then,

$$
\begin{aligned}
\gamma + 2\Gamma \geq \; \frac{(\gamma+2\Gamma)[\gamma(M-2)+2\Gamma]}{\sqrt{\gamma^2(M-2)^2+4\gamma\Gamma M+4\Gamma^2}} \; &\geq \; (\gamma+2\Gamma) - \frac{2\gamma(\gamma+2\Gamma)}{\gamma M+2\Gamma} \\
2\gamma \geq \; \gamma - 2\Gamma + \frac{(\gamma+2\Gamma)[\gamma(M-2)+2\Gamma]}{\sqrt{\gamma^2(M-2)^2+4\gamma\Gamma M+4\Gamma^2}} \; &\geq \; 2\gamma\left[1 - \frac{\gamma+2\Gamma}{\gamma M+2\Gamma}\right] \; > \; 0
\end{aligned}
$$

for $M > 1$. Thus, to show that λ^* is increasing, it remains to

$$
\gamma^2(M-2) + (\gamma+2\Gamma)\sqrt{4\Gamma^2+\gamma^2(M-2)^2+4\gamma\Gamma M} - 2\Gamma(\gamma+2\Gamma+\gamma M)
$$

evaluated at $M =$ is non-negative. When $M = 2$, the expression above is

$$
2\Big[(\gamma+2\Gamma)\sqrt{\Gamma^2+2\gamma\Gamma} \; - \; \Gamma(2\Gamma+3\gamma)\Big] \; > \; 0
$$

for $\gamma, \Gamma > 0$.

Q.E.D.

Proof of Proposition 7. Recall that

$$
\begin{aligned}
\boldsymbol{\delta} &= \; \frac{[2+\gamma\Lambda]\mathbf{a} \; - \; \gamma A\boldsymbol{\lambda}}{4+2\gamma\Lambda}; & \boldsymbol{\eta} &= \; -\frac{\gamma\boldsymbol{\lambda}}{2+\gamma\Lambda}; \\
\partial_{a_i}\delta_j &= \; \frac{\mathbb{1}_i(j)[2+\gamma\Lambda] \; - \; \gamma\lambda_j}{4+2\gamma\Lambda}; & \partial_{a_i}\boldsymbol{\delta} &= \; \frac{1}{4+2\gamma\Lambda}\Big[(2+\gamma\Lambda)\boldsymbol{e}_i - \gamma\boldsymbol{\lambda}\Big],
\end{aligned}
$$

where $\boldsymbol{e}_i$ is the ith coordinate vector. Consider the first FOC:

$$
\begin{aligned}
0 = \; & n(\partial_{a_i}\delta_i)\Big[\Gamma_i x_i - \lambda_i^{-1}a_i\Big] + n\lambda_i^{-1}\Big[(2-n\Gamma_i\lambda_i)(\partial_{a_i}\delta_i) - 1\Big](\delta_i + \eta_i\bar{\theta}) \\
& - \; n(\Gamma_i+2\beta_i)(\Gamma_i+\beta_i)^{-2}\Big(\Psi_i \cdot \partial_{a_i}\boldsymbol{\delta}\Big)\Big(\Psi_i \cdot (x \; - \; n\boldsymbol{\delta} - n\bar{\theta}\boldsymbol{\eta})\Big).
\end{aligned}
$$

We cancel n and multiply by λ_i and get

$$
\begin{aligned}
0 =\ & (\partial_{a_i}\delta_i)\Big[\lambda_i\Gamma_i x_i - a_i\Big] + \Big[(2-n\Gamma_i\lambda_i)(\partial_{a_i}\delta_i) - 1\Big](\delta_i + \eta_i\bar{\theta}) \\
& -\ \lambda_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\Big(\Psi_i\cdot(x\ -\ n\boldsymbol{\delta} - n\bar{\theta}\boldsymbol{\eta})\Big) \\
=\ & (\partial_{a_i}\delta_i)\Big[\lambda_i\Gamma_i x_i - a_i\Big] + \Big[(2-n\Gamma_i\lambda_i)(\partial_{a_i}\delta_i) - 1\Big]\frac{[2+\gamma\Lambda]a_i\ -\ \gamma(A+2\bar{\theta})\lambda_i}{4+2\gamma\Lambda} \\
& -\ \lambda_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\left(\Psi_i\cdot x\ -\ n\Psi_i\cdot\frac{[2+\gamma\Lambda]\mathbf{a}\ -\ \gamma(A+2\bar{\theta})\boldsymbol{\lambda}}{4+2\gamma\Lambda}\right) \\
=\ & (\partial_{a_i}\delta_i)\Big[\lambda_i\Gamma_i x_i - a_i\Big] + \Big[(2-n\Gamma_i\lambda_i)(\partial_{a_i}\delta_i) - 1\Big]\frac{[2+\gamma\Lambda]a_i\ -\ \gamma(A+2\bar{\theta})\lambda_i}{4+2\gamma\Lambda} \\
& -\ \lambda_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\left(\tilde{\chi}\ -\ \frac{n}{2}\tilde{A}\ +\ n\frac{\gamma(A+2\bar{\theta})}{4+2\gamma\Lambda}\tilde{\Lambda}\ -\ \Gamma_i x_i\ +\ \frac{n\Gamma_i}{2}a_i\ -\ \frac{\gamma(A+2\bar{\theta})n\Gamma_i}{4+2\gamma\Lambda}\lambda_i\right) \\
=\ & -\frac{1}{2}\left[n\Gamma_i\lambda_i(\partial_{a_i}\delta_i)\ +\ 1\ +\ n\lambda_i\Gamma_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\right]a_i \\
+\ & \lambda_i\Gamma_i\left[(\partial_{a_i}\delta_i)\ +\ \frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\right]x_i\ -\ \lambda_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\tilde{\chi} \\
& +\ \frac{n}{2}\lambda_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\tilde{A} \\
& -\ \lambda_i\left[n\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\Big(\tilde{\Lambda}\ -\ \Gamma_i\lambda_i\Big)\ +\ (2-n\Gamma_i\lambda_i)(\partial_{a_i}\delta_i)\ -\ 1\right]\frac{\gamma(A+2\bar{\theta})}{4+2\gamma\Lambda} \\
=\ & -\frac{1}{2}\left[1\ +\ nc_{0x_i}\right]a_i\ +\ c_{0x_i}x_i\ -\ \lambda_i\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\left(\tilde{\chi}\ -\ \frac{n}{2}\tilde{A}\right) \\
& -\ \frac{\gamma\lambda_i}{4+2\gamma\Lambda}\left[n\frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\tilde{\Lambda}\ -\ nc_{0x_i}\ +\ 2(\partial_{a_i}\delta_i)\ -\ 1\right](A+2\bar{\theta}),
\end{aligned}
$$

where

$$
\begin{aligned}
c_{0x_i} =\ & \lambda_i\Gamma_i\left[(\partial_{a_i}\delta_i)\ +\ \frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\Big(\Psi_i\cdot\partial_{a_i}\boldsymbol{\delta}\Big)\right] \\
=\ & \lambda_i\Gamma_i\left[\frac{[2+\gamma\Lambda]\ -\ \gamma\lambda_i}{4+2\gamma\Lambda}\ +\ \frac{\Gamma_i+2\beta_i}{(\Gamma_i+\beta_i)^2}\left(\frac{\Gamma_i}{2}\left(\frac{1}{\mathcal{B}(\Gamma_i+\beta_i)} - 1\right)\ -\ \frac{\gamma(\tilde{\Lambda}-\Gamma_i\lambda_i)}{4+2\gamma\Lambda}\right)\right].
\end{aligned}
$$

It follows that

$$a_i = \frac{2c_{0x_i}}{1 \ + \ nc_{0x_i}} x_i \ - \ \frac{2\lambda_i}{1 \ + \ nc_{0x_i}} \frac{\Gamma_i + 2\beta_i}{(\Gamma_i + \beta_i)^2} \Big(\Psi_i \cdot \partial_{a_i} \boldsymbol{\delta} \Big) \left(\tilde{\chi} \ - \ \frac{n}{2} \tilde{A} \right)$$
$$- \ \frac{2}{1 \ + \ nc_{0x_i}} \frac{\gamma \lambda_i}{4 + 2\gamma\Lambda} \left[n \frac{\Gamma_i + 2\beta_i}{(\Gamma_i + \beta_i)^2} \Big(\Psi_i \cdot \partial_{a_i} \boldsymbol{\delta} \Big) \tilde{\Lambda} \ - \ nc_{0x_i} \ + \ 2(\partial_{a_i} \delta_i) \ - \ 1 \right] (A + 2\bar{\theta}) \, ,$$

We now write the Taylor series expansion of this FOC. We shall use the following lemma:

Lemma 12 *When the dispersion of $\Gamma_{i,t}$ is sufficiently small, we have*

$$\frac{\Gamma_{i,t}}{\mathcal{B}_t(\Gamma_{i,t} + \beta_{i,t})} \approx \frac{\Gamma_t^*}{M} + \frac{1}{M^2 - 2M + 2} (\Gamma_{i,t} - \Gamma_t^*)$$
$$\frac{1}{\Gamma_{i,t} + \beta_{i,t}} \approx \frac{M-2}{\Gamma_t^*(M-1)} \ - \ \frac{(M-2)^2}{(M^2 - 2M + 2)(\Gamma_t^*)^2} (\Gamma_{i,t} - \Gamma_t^*)$$
$$\frac{\beta_{i,t}}{\Gamma_{i,t} + \beta_{i,t}} \approx \frac{1}{M-1} \ - \ \frac{M(M-2)}{(M^2 - 2M + 2)(M-1)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \, .$$

It thus follows that

$$\frac{\Gamma_{i,t}}{\Gamma_{i,t} + \beta_{i,t}} \approx \frac{M-2}{M-1} \ + \ \frac{M(M-2)}{(M^2 - 2M + 2)(M-1)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*)$$
$$\frac{\Gamma_{i,t} + 2\beta_{i,t}}{(\Gamma_{i,t} + \beta_{i,t})^2} = \frac{\Gamma_{i,t}}{(\Gamma_{i,t} + \beta_{i,t})^2} \ + \ \frac{2\beta_{i,t}}{(\Gamma_{i,t} + \beta_{i,t})^2}$$
$$\approx \frac{M-2}{M-1} \left[1 \ + \ \frac{M}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \right] \frac{M-2}{\Gamma_t^*(M-1)} \left[1 \ - \ \frac{(M-1)(M-2)}{(M^2 - 2M + 2)\Gamma^*} (\Gamma_{i,t} - \Gamma_t^*) \right]$$
$$+ \ \frac{2}{M-1} \left[1 \ - \ \frac{M(M-2)}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \right] \frac{M-2}{\Gamma_t^*(M-1)} \left[1 \ - \ \frac{(M-1)(M-2)}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \right]$$
$$= \frac{(M-2)^2}{\Gamma_t^*(M-1)^2} \left[1 \ + \ \frac{M}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \ - \ \frac{(M-1)(M-2)}{(M^2 - 2M + 2)\Gamma^*} (\Gamma_{i,t} - \Gamma_t^*) \right]$$
$$+ \ \frac{2(M-2)}{\Gamma_t^*(M-1)^2} \left[1 \ - \ \frac{M(M-2)}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \ - \ \frac{(M-1)(M-2)}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \right]$$
$$= \frac{M(M-2)}{\Gamma_t^*(M-1)^2} \left[1 \ - \ \frac{M(M-2)}{(M^2 - 2M + 2)\Gamma_t^*} (\Gamma_{i,t} - \Gamma_t^*) \right] \, .$$

Moreover,

$$\lambda_i\Gamma_i \;\approx\; \lambda^*\Gamma^* \;+\; [\Phi^\Gamma(\Gamma^*)\Gamma^* + \lambda^*](\Gamma_i - \Gamma^*) \qquad \text{and} \qquad \tilde{\Lambda} \;\approx\; \lambda^*\Gamma^*$$

and

$$\tilde{A} = \sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j+\beta_j)} a_j \;\approx\; \frac{\Gamma^*}{M} A \;+\; \frac{1}{M^2-2M+2} A^\Gamma$$

$$\tilde{\chi} = \sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j+\beta_j)} x_j \;\approx\; \frac{\Gamma^*}{M} X \;+\; \frac{1}{M^2-2M+2} X_{mismatch}\,,$$

where

$$A^\Gamma \;=\; \sum_{j=1}^{M}(\Gamma_j - \Gamma^*)a_j \qquad \text{and} \qquad X_{mismatch} \;=\; \sum_{j=1}^{M}(\Gamma_j - \Gamma^*)x_j$$

Using the lemma above, to first order approximation, we have

$$\begin{aligned}
a_i = \;& [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)]x_i \;-\; [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)]\left(\tilde{\chi} \;-\; \frac{n}{2}\tilde{A}\right) \;-\; [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)](A + 2\bar{\theta}) \\
=\;& [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)]x_i \;-\; [k_{0\theta} + k_{1\theta}(\Gamma_i - \Gamma^*)](A + 2\bar{\theta}) \;-\; [k_{0X} + k_{1X}(\Gamma_i - \Gamma^*)]\frac{\Gamma^*}{M}\left(X \;-\; \frac{n}{2}A\right) \\
&-\; [k_{0X} + k_{1X}(\Gamma_i - \Gamma^*)]\frac{1}{M^2-2M+2}\left(X_{mismatch} \;-\; \frac{n}{2}A^\Gamma\right) \\
=\;& [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)]x_i \;-\; [k_{0\theta} + k_{1\theta}(\Gamma_i - \Gamma^*)](A + 2\bar{\theta}) \;-\; [k_{0X} + k_{1X}(\Gamma_i - \Gamma^*)]\frac{\Gamma^*}{M}\left(X \;-\; \frac{n}{2}A\right) \\
&-\; \frac{k_{0X}}{M^2-2M+2}\left(X_{mismatch} \;-\; \frac{n}{2}A^\Gamma\right) \\
=\;& [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)]x_i \;-\; [k_{0X} + k_{1X}(\Gamma_i - \Gamma^*)]\frac{\Gamma^*}{M}X \;-\; \frac{k_{0X}}{M^2-2M+2}X_{mismatch} \\
&-\; \left[\frac{2Mk_{0\theta} - nk_{0X}\Gamma^*}{2M} + \frac{2Mk_{1\theta} - nk_{1X}\Gamma^*}{2M}(\Gamma_i - \Gamma^*)\right] A \;+\; \frac{n}{2}\frac{k_{0X}}{M^2-2M+2}A^\Gamma \\
&-\; [k_{0\theta} + k_{1\theta}(\Gamma_i - \Gamma^*)]2\bar{\theta}\,,
\end{aligned}$$

Summing over i we get

$$
\begin{aligned}
A^\Gamma &= \; k_{0x} X_{mismatch} \\
A &= \; k_{0x} X + k_{1x} X_{mismatch} \; - \; k_{0X} \Gamma^* X \; - \; \frac{M k_{0X}}{M^2 - 2M + 2} X_{mismatch} \\
& \quad - \; \frac{2M k_{0\theta} - n k_{0X} \Gamma^*}{2} A \; + \; \frac{n}{2} \frac{M k_{0X}}{M^2 - 2M + 2} A^\Gamma \; - \; 2M k_{0\theta} \bar{\theta} \, .
\end{aligned}
$$

Thus,

$$
A = \left[1 \; + \; \frac{2M k_{0\theta} - n k_{0X} \Gamma^*}{2} \right]^{-1} \Bigg(\left[k_{0x} - k_{0X} \Gamma^* \right] X \; + \; \left[k_{1x} - \frac{(2 - n k_{0x}) M k_{0X}}{2(M^2 - 2M + 2)} \right] X_{mismatch} \; - \; 2M k
$$

Substituting into the equation for a_i we obtain

$$
\begin{aligned}
a_i = & \; [k_{0x} + k_{1x}(\Gamma_i - \Gamma^*)] x_i \; - \; [k_{0X} + k_{1X}(\Gamma_i - \Gamma^*)] \frac{\Gamma^*}{M} X \\
& - \; \left[\frac{2M k_{0\theta} - n k_{0X} \Gamma^*}{M} + \frac{2M k_{1\theta} - n k_{1X} \Gamma^*}{M} (\Gamma_i - \Gamma^*) \right] \frac{k_{0x} - k_{0X} \Gamma^*}{2 + 2M k_{0\theta} - n k_{0X} \Gamma^*} X \\
& - \; \left(1 \; - \; \frac{n}{2} k_{0x} \right) \frac{k_{0X}}{M^2 - 2M + 2} X_{mismatch} \\
& - \; \left[\frac{2M k_{0\theta} - n k_{0X} \Gamma^*}{M} + \frac{2M k_{1\theta} - n k_{1X} \Gamma^*}{M} (\Gamma_i - \Gamma^*) \right] \frac{k_{1x} - \frac{(2 - n k_{0x}) M k_{0X}}{2(M^2 - 2M + 2)}}{2 + 2M k_{0\theta} - n k_{0X} \Gamma^*} X_{mismatch} \\
& + \; \left[\frac{2M k_{0\theta} - n k_{0X} \Gamma^*}{M} + \frac{2M k_{1\theta} - n k_{1X} \Gamma^*}{M} (\Gamma_i - \Gamma^*) \right] \frac{2M k_{0\theta}}{2 + 2M k_{0\theta} - n k_{0X} \Gamma^*} \bar{\theta} \\
& - \; 2[k_{0\theta} + k_{1\theta}(\Gamma_i - \Gamma^*)] \bar{\theta} \, .
\end{aligned}
$$

It follows that

$$
\begin{aligned}
a_i = & \; \Phi_0^x \, x_i \; + \; \Phi_0^X \, X \; + \; \Phi_0^\Gamma \, X_{mismatch} \; + \; \Phi_0^\theta \, \bar{\theta} \; + \; (\Gamma_i - \Gamma^*) [\Phi_1^x \, x_i + \Phi_1^X \, X + \Phi_1^\theta \, \bar{\theta}] \\
& + \; O((\|\Gamma - \Gamma^*\|^2 + \|x\|^2 + \bar{\theta}^2)^{3/2}) \, ,
\end{aligned}
$$

with

$$\Phi_0^x = \; k_{0x}; \qquad\qquad \Phi_1^x = \; k_{1x}$$

$$\Phi_0^X = \; - \left[\frac{2k_{0X}\Gamma^* \; + \; k_{0x}\left(2Mk_{0\theta} - nk_{0X}\Gamma^*\right)}{2 + 2Mk_{0\theta} - nk_{0X}\Gamma^*}\right]\frac{1}{M};$$

$$\Phi_1^X = \; - \left[\frac{k_{1X}\Gamma^*(2 + 2Mk_{0\theta}) \; + \; 2Mk_{1\theta}(k_{0x} - k_{0X}\Gamma^*) \; - \; nk_{1X}k_{0x}\Gamma^*}{2 + 2Mk_{0\theta} - nk_{0X}\Gamma^*}\right]\frac{1}{M}$$

$$\Phi_0^\Gamma = \; - \left[\frac{k_{0X}(2 - nk_{0x})}{M^2 - 2M + 2} \; + \; \frac{k_{1x}(2Mk_{0\theta} - nk_{0X}\Gamma^*)}{M}\right]\frac{1}{2 + 2Mk_{0\theta} - nk_{0X}\Gamma^*}$$

$$\Phi_0^\theta = \; \frac{-4k_{0\theta}}{2 + 2Mk_{0\theta} - nk_{0X}\Gamma^*}$$

$$\Phi_1^\theta = \; - \frac{2\left[2k_{1\theta} \; - \; n\left(k_{0X}k_{1\theta} \; - \; k_{0\theta}k_{1X}\right)\Gamma^*\right]}{2 + 2Mk_{0\theta} - nk_{0X}\Gamma^*}\,.$$

Q.E.D.

Proof of Proposition 8. The proof is similar to that of Proposition 7. The FOC is

$$\begin{aligned}0 = \; & -\frac{1}{2}\left[1 \; + \; nc_{0x_i}\right]a_i \; + \; c_{0x_i}x_i \; - \; \lambda_i\frac{\Gamma_i + 2\beta_i}{(\Gamma_i + \beta_i)^2}\Big(\Psi_i \cdot \partial_{a_i}\boldsymbol{\delta}\Big)\left(\tilde{\chi} \; - \; \frac{n}{2}\tilde{A}\right) \\ & - \; \frac{\gamma\lambda_i}{4 + 2\gamma\Lambda}\left[n\frac{\Gamma_i + 2\beta_i}{(\Gamma_i + \beta_i)^2}\Big(\Psi_i \cdot \partial_{a_i}\boldsymbol{\delta}\Big)\tilde{\Lambda} \; - \; nc_{0x_i} \; + \; 2(\partial_{a_i}\delta_i) \; - \; 1\right](A + 2\bar{\theta})\,.\end{aligned}$$

It follows that

$$\begin{aligned}x_i = \; & \frac{1 \; + \; nc_{0x_i}}{2c_{0x_i}}a_i \; + \; \frac{\lambda_i}{c_{0x_i}}\frac{\Gamma_i + 2\beta_i}{(\Gamma_i + \beta_i)^2}\Big(\Psi_i \cdot \partial_{a_i}\boldsymbol{\delta}\Big)\left(\tilde{\chi} \; - \; \frac{n}{2}\tilde{A}\right) \\ & + \; \frac{1}{c_{0x_i}}\frac{\gamma\lambda_i}{4 + 2\gamma\Lambda}\left[n\frac{\Gamma_i + 2\beta_i}{(\Gamma_i + \beta_i)^2}\Big(\Psi_i \cdot \partial_{a_i}\boldsymbol{\delta}\Big)\tilde{\Lambda} \; - \; nc_{0x_i} \; + \; 2(\partial_{a_i}\delta_i) \; - \; 1\right](A + 2\bar{\theta}) \\ \approx \; & [h_{0a} + h_{1a}(\lambda_i - \lambda^*)]a_i \; + \; [h_{0X} + h_{1X}(\lambda_i - \lambda^*)]\left(\tilde{\chi} \; - \; \frac{n}{2}\tilde{A}\right) \; + \; [h_{0\theta} + h_{1\theta}(\lambda_i - \lambda^*)](A + 2\bar{\theta}) \\ = \; & [h_{0a} + h_{1a}(\lambda_i - \lambda^*)]a_i \; + \; \frac{\Gamma^*[h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{M}\left(X \; - \; \frac{n}{2}A\right) \; + \; [h_{0\theta} + h_{1\theta}(\lambda_i - \lambda^*)](A + 2\bar{\theta}) \\ & + \; \frac{\varphi^\lambda[h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{M^2 - 2M + 2}\left(X^\lambda \; - \; \frac{n}{2}A^\lambda\right)\,,\end{aligned}$$

where

$$A^{\lambda} \;=\; \sum_{j=1}^{M} (\lambda_j - \lambda^*) a_j \qquad \text{and} \qquad X^{\lambda} \;=\; \sum_{j=1}^{M} (\lambda_j - \lambda^*) x_j \,,$$

and we used

$$\begin{aligned}
\tilde{A} &= \sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j + \beta_j)} a_j \;\approx\; \frac{\Gamma^*}{M} A \;+\; \frac{\varphi^{\lambda}}{M^2 - 2M + 2} A^{\Gamma} \\
\tilde{\chi} &= \sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j + \beta_j)} x_j \;\approx\; \frac{\Gamma^*}{M} X \;+\; \frac{\varphi^{\lambda}}{M^2 - 2M + 2} X_{mismatch} \,.
\end{aligned}$$

Summing over i yield

$$\begin{aligned}
X^{\lambda} &= \; h_{0a} A^{\lambda} \\
X &= \; h_{0a} A + h_{1a} A^{\lambda} \;+\; \Gamma^* h_{0X} \left(X \;-\; \frac{n}{2} A \right) \;+\; M h_{0\theta} (A + 2\bar{\theta}) \;+\; \frac{M h_{0X} \varphi^{\lambda}}{M^2 - 2M + 2} \left(X^{\lambda} \;-\; \frac{n}{2} A^{\lambda} \right) \,,
\end{aligned}$$

implying that

$$X = \; \frac{h_{0a} - \frac{n}{2} \Gamma^* h_{0X} + M h_{0\theta}}{1 - \Gamma^* h_{0X}} A \;+\; \frac{2 M h_{0\theta}}{1 - \Gamma^* h_{0X}} \bar{\theta} \;+\; \frac{h_{1a} \;+\; \frac{M h_{0X} \varphi^{\lambda}}{M^2 - 2M + 2} \left(h_{0a} \;-\; \frac{n}{2} \right)}{1 - \Gamma^* h_{0X}} A^{\lambda} \,.$$

Substituting into the equation for x_i we obtain

$$\begin{aligned}
x_i = \; & [h_{0a} + h_{1a}(\lambda_i - \lambda^*)] a_i \;+\; \frac{\Gamma^* [h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{M} \frac{h_{0a} - \frac{n}{2} \Gamma^* h_{0X} + M h_{0\theta}}{1 - \Gamma^* h_{0X}} A \\
& -\; \frac{n}{2} \frac{\Gamma^* [h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{M} A \;+\; [h_{0\theta} + h_{1\theta}(\lambda_i - \lambda^*)] A \;+\; \left(h_{0a} - \frac{n}{2} \right) \frac{\varphi^{\lambda} [h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{M^2 - 2M + 2} A^{\lambda} \\
& +\; \frac{\Gamma^* [h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{M} \frac{h_{1a} \;+\; \frac{M h_{0X} \varphi^{\lambda}}{M^2 - 2M + 2} \left(h_{0a} \;-\; \frac{n}{2} \right)}{1 - \Gamma^* h_{0X}} A^{\lambda} \\
& +\; 2 [h_{0\theta} + h_{1\theta}(\lambda_i - \lambda^*)] \bar{\theta} \;+\; \frac{2 h_{0\theta} \Gamma^* [h_{0X} + h_{1X}(\lambda_i - \lambda^*)]}{1 - \Gamma^* h_{0X}} \bar{\theta} \,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\varphi_0^a =&\ h_{0a} \\
\varphi_0^A =&\ h_{0\theta} \ + \ \frac{\Gamma^* h_{0X}}{M} \frac{h_{0a} - \frac{n}{2}\Gamma^* h_{0X} + M h_{0\theta}}{1 - \Gamma^* h_{0X}} \ - \ \frac{n}{2}\frac{\Gamma^* h_{0X}}{M} \\
=&\ h_{0\theta} \ + \ \frac{\Gamma^* h_{0X}}{M} \frac{1}{1 - \Gamma^* h_{0X}} \left[h_{0a} \ + \ M h_{0\theta} \ - \ \frac{n}{2}\right] \\
\varphi_{0,\lambda}^A =&\ \left(h_{0a} - \frac{n}{2}\right) \frac{\varphi^\lambda h_{0X}}{M^2 - 2M + 2} \ + \ \frac{\Gamma^* h_{0X}}{M} \frac{h_{1a} + \frac{M h_{0X}\varphi^\lambda}{M^2-2M+2}\left(h_{0a} - \frac{n}{2}\right)}{1 - \Gamma^* h_{0X}} \\
\varphi_0^\theta =&\ 2h_{0\theta}\left[1 \ + \ \frac{\Gamma^* h_{0X}}{1 - \Gamma^* h_{0X}}\right] \ = \ \frac{2h_{0\theta}}{1 - \Gamma^* h_{0X}} \\
\varphi_1^a =&\ h_{1a} \\
\varphi_1^A =&\ h_{1\theta} \ + \ \frac{\Gamma^* h_{1X}}{M} \frac{h_{0a} - \frac{n}{2}\Gamma^* h_{0X} + M h_{0\theta}}{1 - \Gamma^* h_{0X}} \ - \ \frac{n}{2}\frac{\Gamma^* h_{1X}}{M} \\
=&\ h_{1\theta} \ + \ \frac{\Gamma^* h_{1X}}{M} \frac{1}{1 - \Gamma^* h_{0X}} \left[h_{0a} \ + \ M h_{0\theta} \ - \ \frac{n}{2}\right] \\
\varphi_{1,\lambda}^A =&\ \left(h_{0a} - \frac{n}{2}\right) \frac{\varphi^\lambda h_{1X}}{M^2 - 2M + 2} \ + \ \frac{\Gamma^* h_{1X}}{M} \frac{h_{1a} + \frac{M h_{0X}\varphi^\lambda}{M^2-2M+2}\left(h_{0a} - \frac{n}{2}\right)}{1 - \Gamma^* h_{0X}} \\
=&\ \frac{1}{1 - \Gamma^* h_{0X}} \left[\left(h_{0a} - \frac{n}{2}\right) \frac{\varphi^\lambda h_{1X}}{M^2 - 2M + 2} \ + \ \frac{\Gamma^* h_{1X} h_{1a}}{M}\right] \\
\varphi_1^\theta =&\ 2h_{1\theta} \ + \ \frac{2h_{0\theta}\Gamma^* h_{1X}}{1 - \Gamma^* h_{0X}} \,.
\end{aligned}$$

Q.E.D.

Lemma 13 *After the D2C trading round, dealer inventories become*

$$\begin{aligned}
\chi_i \approx&\ -(\varphi_0^a - 0.5n)\lambda^*\alpha_i \ - \ (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*)M\lambda^*\bar{\alpha} \ - \ M\lambda^*[(\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*) + \varphi_{0,\lambda}^A\lambda^*]\hat{\alpha} \\
&+ \ \lambda^*[\varphi_0^a + M\varphi_0^A + n\eta_0^*]d \ + \ \varphi_0^\theta\bar{\theta} \ - \ \gamma\eta_0^*\lambda^*\Theta \\
&+ (\lambda_i - \lambda^*)\left[-(\varphi_0^a + \varphi_1^a\lambda^* - 0.5n)\alpha_i + M\lambda^*(0.5n\gamma\eta_0^* - \varphi_1^A)\bar{\alpha}\right] \\
&+ \ (\lambda_i - \lambda^*)[(\varphi_0^a + n\eta_0^* + \lambda^*\varphi_1^a + M\lambda^*\varphi_1^A)d - \gamma\eta_0^*\Theta + \varphi_1^\theta\bar{\theta}]\,,
\end{aligned}$$

where

$$\bar{\alpha} \;=\; \frac{1}{M}\sum_i \alpha_i; \quad \hat{\alpha} \;=\; \frac{1}{\Lambda}\sum_i (\lambda_i - \lambda^*)\alpha_i; \quad \text{and} \quad \eta_0^* \;=\; -\frac{1}{2+\gamma\Lambda}\,.$$

Proof of Lemma 13.

$$\begin{aligned}
\tilde{\chi}_i =&\; x_i \;-\; \sum_j (\delta_i \;+\; \eta_i\theta_{j,t-}) \;=\; x_i \;-\; 0.5na_i \;-\; \eta_i(\Theta_t + 0.5nA) \\
=&\; \varphi_0^a a_i \;+\; \varphi_0^A A \;+\; \varphi_{0,\lambda}^A A^\lambda \;+\; \varphi_0^\theta \bar{\theta} \;+\; (\lambda_i - \lambda^*)[\varphi_1^a a_i + \varphi_1^A A + \varphi_{1,\lambda}^A A^\lambda + \varphi_1^\theta \bar{\theta}] \\
&- \; 0.5na_i \;-\; \eta_i(\Theta + 0.5nA) \\
=&\; (\varphi_0^a - 0.5n)a_i \;+\; (\varphi_0^A - 0.5n\eta_i)A \;+\; \varphi_{0,\lambda}^A A^\lambda \;+\; \varphi_0^\theta \bar{\theta} \;-\; \eta_i\Theta \;+\; (\lambda_i - \lambda^*)[\varphi_1^a a_i + \varphi_1^A A + \varphi_1^\theta \bar{\theta}]\,.
\end{aligned}$$

For simplicity, we assume that the total customer inventory shocks $\Theta_t = \sum_j \theta_{j,t}$ are i.i.d. over time. We rerwite both a_i and A in terms of mid-prices:

$$\begin{aligned}
a_i =&\; \lambda_i(d - \alpha_i) \\
A =&\; \sum_j a_j \;=\; d\sum_j \lambda_j \;-\; \sum_j \lambda_j\alpha_j \;=\; d\Lambda \;-\; \sum_j \lambda_j\alpha_j \\
=&\; M\lambda^*\left[d \;-\; (\hat{\alpha} + \bar{\alpha})\right] \\
A^\lambda =&\; \sum_j (\lambda_j - \lambda^*)a_j \;=\; d\sum_j (\lambda_j - \lambda^*)\lambda_j \;-\; \sum_j (\lambda_j - \lambda^*)\lambda_j\alpha_j \;\approx\; -\,\lambda^*\sum_j (\lambda_j - \lambda^*)\alpha_j \\
=&\; -\,M(\lambda^*)^2\hat{\alpha}\,.
\end{aligned}$$

Thus,

$$
\begin{aligned}
\tilde{\chi}_i =\; & (\varphi_0^a - 0.5n)\lambda_i(d-\alpha_i) \;+\; (\varphi_0^A - 0.5n\eta_i)M\lambda^*\left[d \;-\; (\hat{\alpha}+\bar{\alpha})\right] \;-\; \varphi_{0,\lambda}^A M(\lambda^*)^2\hat{\alpha} \\
& +\; \varphi_0^\theta\bar{\theta} \;-\; \eta_i\Theta \;+\; (\lambda_i-\lambda^*)[\varphi_1^a\lambda_i(d-\alpha_i) + \varphi_1^A M\lambda^*\left[d \;-\; (\hat{\alpha}+\bar{\alpha})\right] + \varphi_1^\theta\bar{\theta}] \\
=\; & -(\varphi_0^a - 0.5n)\lambda_i\alpha_i \;-\; (\varphi_0^A - 0.5n\gamma\lambda_i\eta_0^*)M\lambda^*\bar{\alpha} \;-\; M\lambda^*[(\varphi_0^A - 0.5n\gamma\lambda_i\eta_0^*) + \varphi_{0,\lambda}^A\lambda^*]\hat{\alpha} \\
& +\; [(\varphi_0^a - 0.5n)\lambda_i \;+\; (\varphi_0^A - 0.5n\gamma\lambda_i\eta_0^*)M\lambda^*]d \;+\; \varphi_0^\theta\bar{\theta} \;-\; \gamma\lambda_i\eta_0^*\Theta \\
& +\; (\lambda_i-\lambda^*)[-\varphi_1^a\lambda_i\alpha_i - \varphi_1^A M\lambda^*(\hat{\alpha}+\bar{\alpha}) + (\varphi_1^a\lambda_i + \varphi_1^A M\lambda^*)d + \varphi_1^\theta\bar{\theta}] \\
\approx\; & -(\varphi_0^a - 0.5n)\lambda^*\alpha_i \;-\; (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*)M\lambda^*\bar{\alpha} \;-\; M\lambda^*[(\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*) + \varphi_{0,\lambda}^A\lambda^*]\hat{\alpha} \\
& +\; [(\varphi_0^a - 0.5n)\lambda^* \;+\; (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*)M\lambda^*]d \;+\; \varphi_0^\theta\bar{\theta} \;-\; \gamma\eta_0^*\lambda^*\Theta \\
& -(\varphi_0^a - 0.5n)(\lambda_i-\lambda^*)\alpha_i \;+\; 0.5n\gamma\eta_0^*(\lambda_i-\lambda^*)M\lambda^*\bar{\alpha} \\
& +\; [(\varphi_0^a - 0.5n)(\lambda_i-\lambda^*) \;-\; 0.5n\gamma\eta_0^*(\lambda_i-\lambda^*)M\lambda^*]d \;-\; \gamma\eta_0^*(\lambda_i-\lambda^*)\Theta \\
& +\; (\lambda_i-\lambda^*)[-\varphi_1^a\lambda^*\alpha_i - \varphi_1^A M\lambda^*\bar{\alpha} + \lambda^*(\varphi_1^a + M\varphi_1^A)d + \varphi_1^\theta\bar{\theta}] \\
\approx\; & -(\varphi_0^a - 0.5n)\lambda^*\alpha_i \;-\; (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*)M\lambda^*\bar{\alpha} \;-\; M\lambda^*[(\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*) + \varphi_{0,\lambda}^A\lambda^*]\hat{\alpha} \\
& +\; [(\varphi_0^a - 0.5n)\lambda^* \;+\; (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*)M\lambda^*]d \;+\; \varphi_0^\theta\bar{\theta} \;-\; \gamma\eta_0^*\lambda^*\Theta \\
& +(\lambda_i-\lambda^*)\left[-(\varphi_0^a - 0.5n)\alpha_i \;+\; 0.5nM\gamma\eta_0^*\lambda^*\bar{\alpha}\right] \\
& +\; (\lambda_i-\lambda^*)[(\varphi_0^a - 0.5n) \;-\; 0.5nM\gamma\eta_0^*\lambda^*]d \;-\; (\lambda_i-\lambda^*)\gamma\eta_0^*\Theta \\
& +\; (\lambda_i-\lambda^*)[-\varphi_1^a\lambda^*\alpha_i - \varphi_1^A M\lambda^*\bar{\alpha} + \lambda^*(\varphi_1^a + M\varphi_1^A)d + \varphi_1^\theta\bar{\theta}]\,,
\end{aligned}
$$

where we used the definition

$$
\eta_i \;=\; -\frac{\gamma\lambda_i}{2+\gamma\Lambda} \;=\; \gamma\eta_0^*\lambda_i \qquad \Longrightarrow \qquad 1 \;+\; M\gamma\eta_0^*\lambda^* \;=\; -\,2\eta_0^*\,.
$$

Thus,

$$\begin{aligned}
\tilde{\chi}_i \approx\; & -(\varphi_0^a - 0.5n)\lambda^* \alpha_i \;-\; (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*)M\lambda^*\bar{\alpha} \;-\; M\lambda^*[(\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*) + \varphi_{0,\lambda}^A\lambda^*]\hat{\alpha} \\
& +\; \lambda^*[\varphi_0^a + M\varphi_0^A + n\eta_0^*]d \;+\; \varphi_0^\theta\bar{\theta} \;-\; \gamma\eta_0^*\lambda^*\Theta \\
& + (\lambda_i - \lambda^*)\,[-(\varphi_0^a - 0.5n)\alpha_i + 0.5nM\gamma\eta_0^*\lambda^*\bar{\alpha}] \\
& +\; (\lambda_i - \lambda^*)[(\varphi_0^a + n\eta_0^*)d - \gamma\eta_0^*\Theta] \\
& +\; (\lambda_i - \lambda^*)[-\varphi_1^a\lambda^*\alpha_i - \varphi_1^A M\lambda^*\bar{\alpha} + \lambda^*(\varphi_1^a + M\varphi_1^A)d + \varphi_1^\theta\bar{\theta}]\,,
\end{aligned}$$

Q.E.D.

Proof of Proposition 9. We have

$$P^{D2D} \;=\; d \;-\; \widehat{X} \qquad \text{and} \qquad \widehat{X} \;=\; \Psi \cdot \tilde{\chi}\,.$$

To first order,

$$\begin{aligned}
\frac{\Gamma_i}{\mathcal{B}(\Gamma_i + \beta_i)} &\approx\; \frac{\Gamma^*}{M} \;+\; \frac{\varphi^\lambda}{M^2 - 2M + 2}(\lambda_i - \lambda^*) \\
\sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j + \beta_j)} &\approx\; \Gamma^* \\
\sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j + \beta_j)}\alpha_j &\approx\; \Gamma^*\bar{\alpha} \;+\; \frac{M\varphi^\lambda\lambda^*}{M^2 - 2M + 2}\hat{\alpha} \\
\sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j + \beta_j)}(\lambda_j - \lambda^*) &\approx\; 0 \\
\sum_{j=1}^{M} \frac{\Gamma_j}{\mathcal{B}(\Gamma_j + \beta_j)}\alpha_j(\lambda_j - \lambda^*) &\approx\; \lambda^*\Gamma^*\hat{\alpha}\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
\Psi \cdot \tilde{\chi} = \; & -(\varphi_0^a - 0.5n)\lambda^* \left[\Gamma^* \bar{\alpha} \; + \; \frac{M\varphi^\lambda \lambda^*}{M^2 - 2M + 2} \hat{\alpha} \right] \\
& - \; (\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*) M\lambda^*\Gamma^*\bar{\alpha} \; - \; M\lambda^*\Gamma^*[(\varphi_0^A - 0.5n\gamma\eta_0^*\lambda^*) + \varphi_{0,\lambda}^A \lambda^*]\hat{\alpha} \\
& + \; \lambda^*\Gamma^*[\varphi_0^a + M\varphi_0^A + n\eta_0^*]d \; + \; \Gamma^*\varphi_0^\theta \bar{\theta} \; - \; \gamma\eta_0^*\lambda^*\Gamma^*\Theta \\
& - (\varphi_0^a + \varphi_1^a \lambda^* - 0.5n)\lambda^*\Gamma^*\hat{\alpha} \\
= \; & -\lambda^*\Gamma^* \left[\varphi_0^a + M\varphi_0^A + n\eta_0^* \right] (\bar{\alpha} + \hat{\alpha}) \; - \; (\lambda^*)^2 \left[\Gamma^*(M\varphi_{0,\lambda}^A + \varphi_1^a) \; + \; \frac{M\varphi^\lambda(\varphi_0^a - 0.5n)}{M^2 - 2M + 2} \right] \hat{\alpha} \\
& + \; \lambda^*\Gamma^*[\varphi_0^a + M\varphi_0^A + n\eta_0^*]d \; + \; \Gamma^*\varphi_0^\theta \bar{\theta} \; - \; \gamma\eta_0^*\lambda^*\Gamma^*\Theta \, .
\end{aligned}$$

Q.E.D.

B Regression Tables: Main Results

Table 2: Forecasting FX rates: Ten Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} \ = \ a_0 \ + \ a_1 \Delta \bar{\alpha}_{t,\ell} \ + \ a_2 \Delta \hat{\alpha}_{t,\ell} \ + \ controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell = 10s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\Delta\hat{\alpha}_{t,\ell}$	0.029***	0.029***	0.029***	0.029***	0.029***	0.029***	0.030***	0.029***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
ProxyLiqShock		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							−0.00000	
							(0.00000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: Forecasting FX Spread: Ten Seconds Ahead.— The table report estimates for the regression

$$\Delta Spread\_D2D_{t+\ell,\ell} \; = \; a_0 \; + \; a_1 \Delta \overline{Spread\_D2C}_{t,\ell} \; + \; a_2 STD\_Spread\_D2C_t \; + \; controls_t \, ,$$

$\Delta X_{t,\ell} = X_{t+\ell} \; - \; X_t$ and $\ell \; = \; 10s$. The variaable $Spread\_D2D$ is the bid-ask spread for the "EUR/USD" exchange rate, from the EBS trading platform. The variable $\overline{Spread\_D2C}$ is the mean spread from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers while the variable STD_Spread_D2C is the standard deviation of bid-ask spread from the same price schedules. The set of control variables are: (1) $ProxyLiqShock$, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_BinHeteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$					
$\Delta\overline{Spread\_D2C}_{t,\ell}$	-0.0002^{***}	-0.0002^{***}	-0.0002^{***}	-0.0002^{***}	-0.0002^{***}	-0.0002^{***}
	(0.00004)	(0.00004)	(0.00004)	(0.00003)	(0.00003)	(0.00003)
$STD\_Spread\_D2C_t$	-0.00000	-0.00000	-0.00000	-0.00000	-0.00000	-0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
OFlow_d2d_Vol$_t$		-0.000			-0.000	
		(0.000)			(0.000)	
OFlow_d2d_Bin$_t$			0.000			0.000
			(0.000)			(0.000)
ProxyLiqShock$_t$				0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)
Constant	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

C Regression Tables: Additional Lags

Table 4: Forecasting FX rates: One Second Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} \;=\; a_0 \;+\; a_1 \Delta\bar{\alpha}_{t,\ell} \;+\; a_2 \Delta\hat{\alpha}_{t,\ell} \;+\; controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell \;=\; 1s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\Delta\bar{\alpha}_{t,\ell}$	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							−0.000	
							(0.000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	0.000**	0.000*	0.000	0.000*	0.000	0.000*	0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 5: Forecasting FX rates: Three Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} \;=\; a_0 \;+\; a_1 \Delta\bar{\alpha}_{t,\ell} \;+\; a_2 \Delta\hat{\alpha}_{t,\ell} \;+\; controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell \;=\; 3s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
$\Delta\hat{\alpha}_{t,\ell}$	0.00001**	0.00001*	0.00001**	0.00001**	0.00001*	0.00001*	0.00001*	0.00001*
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			−0.000		−0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				−0.000		−0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							−0.000***	
							(0.000)	
OF_Vol_24h$_t$								−0.000
								(0.000)
Constant	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 6: Forecasting FX rates: Five Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} \;=\; a_0 \;+\; a_1 \Delta\bar{\alpha}_{t,\ell} \;+\; a_2 \Delta\hat{\alpha}_{t,\ell} \;+\; controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell \;=\; 5s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
$\Delta\hat{\alpha}_{t,\ell}$	0.00001**	0.00001**	0.00001***	0.00001***	0.00001**	0.00001**	0.00001**	0.00001**
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				−0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							−0.000***	
							(0.000)	
OF_Vol_24h$_t$								0.000***
								(0.000)
Constant	0.000	0.000	0.000	0.000	0.000	0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 7: Forecasting FX rates: Fifteen Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} \;=\; a_0 \;+\; a_1 \Delta\bar{\alpha}_{t,\ell} \;+\; a_2 \Delta\hat{\alpha}_{t,\ell} \;+\; controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell = 15s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.015***	0.015***	0.015***	0.015***	0.015***	0.015***	0.016***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\Delta\hat{\alpha}_{t,\ell}$	0.012***	0.012***	0.012***	0.012***	0.012***	0.012***	0.013***	0.012***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			−0.000		−0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				−0.000		−0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							−0.00000	
							(0.00000)	
OF_Vol_24h$_t$								−0.000
								(0.000)
Constant	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Table 8: Forecasting FX rates: Twenty Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} = a_0 + a_1 \Delta\bar{\alpha}_{t,\ell} + a_2 \Delta\hat{\alpha}_{t,\ell} + controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell = 20s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\Delta\hat{\alpha}_{t,\ell}$	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							−0.00000	
							(0.00000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Table 9: Forecasting FX rates: Thirty Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} = a_0 + a_1 \Delta\bar{\alpha}_{t,\ell} + a_2 \Delta\hat{\alpha}_{t,\ell} + controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell = 30s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	0.005*	0.005*	0.005*	0.005*	0.005*	0.005*	0.005*	0.005*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta\hat{\alpha}_{t,\ell}$	0.011**	0.011**	0.011**	0.011**	0.011**	0.011**	0.011**	0.011**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							0.000	
							(0.00000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Table 10: Forecasting FX rates: Forty Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} = a_0 + a_1 \Delta\bar{\alpha}_{t,\ell} + a_2 \Delta\hat{\alpha}_{t,\ell} + controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell = 40s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	−0.0001	0.0001	−0.00004	−0.00005	0.0001	0.0001	−0.0001	0.00005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta\hat{\alpha}_{t,\ell}$	−0.014	−0.014	−0.014	−0.014	−0.014	−0.014	−0.014	−0.014
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							0.00000	
							(0.00000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Table 11: Forecasting FX rates: Fourty Five Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} = a_0 + a_1 \Delta\bar{\alpha}_{t,\ell} + a_2 \Delta\hat{\alpha}_{t,\ell} + controls_t$$

where $\Delta X_{t,\ell} = X_{t+\ell} - X_t$ and $\ell = 45s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+1,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$\Delta\hat{\alpha}_{t,\ell}$	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							0.00000	
							(0.00000)	
OF_Vol_24$_t$								0.000
								(0.000)
Constant	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Table 12: Forecasting FX rates: Fifty Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} \;=\; a_0 \;+\; a_1 \Delta\bar{\alpha}_{t,\ell} \;+\; a_2 \Delta\hat{\alpha}_{t,\ell} \;+\; controls_t$$

where $\Delta X_{t,\ell} = X_t - X_{t-\ell}$ and $\ell = 50s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	−0.004	−0.004	−0.004	−0.004	−0.004	−0.004	−0.004	−0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta\hat{\alpha}_{t,\ell}$	−0.008	−0.008	−0.008	−0.008	−0.008	−0.008	−0.008	−0.008
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
OFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							0.00000	
							(0.00000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

Table 13: Forecasting FX rates: Sixty Seconds Ahead.— The table report estimates for the regression

$$\Delta p^{D2D}_{t+\ell,\ell} = a_0 + a_1 \Delta\bar{\alpha}_{t,\ell} + a_2 \Delta\hat{\alpha}_{t,\ell} + controls_t$$

where $\Delta X_{t,\ell} = X_{t+\ell} - X_t$ and $\ell = 60s$. The dependent variaable p^{D2D}_t is the log mid-price for the "EUR/USD" exchange rate, from the EBS trading platform. The independent variable $\bar{\alpha}_t$ is the average mid-price from the price schedules for the currency pair "EUR/USD" submitted to Swissquote by a set of large dealers. The independent variable $\hat{\alpha}_t$ is our novel measure of liquidity missmatch in the two-tiered FX market. The set of control variables are: (1) ProxyLiqShock, a proxy for the customers' liquidity demand that is the net of all seller-initiated orders and buyer-initiated orders from the Swissquote's clients (i.e., the net selling pressure); (2) OFlow_d2d_Vol, the aggregate daily net order flow; (3) OFlow_d2d_Bin, the aggregate daily binary Buy and Sell executed, where a buy (Sell) order is order marked as +1 (−1). (3) OF_Vol_1h, the aggregate (signed) order flow in the D2D market over the previous 1 hour; (4) OF_Vol_24h, the aggregate (signed) order flow in the D2D market over the previous 24 hours. Heteroscedasticity and autocorrelation robust standard errors are shown in parenthesis.

	$\Delta p^{D2D}_{t+\ell,\ell}$							
$\Delta\bar{\alpha}_{t,\ell}$	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$\Delta\hat{\alpha}_{t,\ell}$	−0.016	−0.016	−0.016	−0.016	−0.016	−0.016	−0.016	−0.016
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
ProxyLiqShock$_t$		−0.000***			−0.000***	−0.000***	−0.000***	−0.000***
		(0.000)			(0.000)	(0.000)	(0.000)	(0.000)
LOFlow_d2d_Vol$_t$			0.000		0.000			
			(0.000)		(0.000)			
OFlow_d2d_Bin$_t$				0.000		0.000		
				(0.000)		(0.000)		
OF_Vol_1h$_t$							0.00000	
							(0.00000)	
OF_Vol_24h$_t$								0.000
								(0.000)
Constant	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)