

Continuous Time Decision-Making in a Partially Decentralized Multiple Dealership Forex Market, and the Equilibrium Exchange Rate

Alexis Derviz¹²

ABSTRACT

The paper introduces a model of bid/ask price formation in an imperfectly centralized forex dealership market in continuous time. The dealers have costly access to best quotes while interpreting signals from the joint dealer-customer order flow and deciding upon their own price quotes and activities in the inter-dealer market. The lowest ask/highest bid price of a foreign currency becomes the single transaction price in equilibrium. Each dealer uses the observed order flow to improve the subjective estimates of relevant aggregate variables, which are the sources of uncertainty. These are: returns on domestic and foreign assets, the equilibrium FX transaction price and the size of the aggregate dealer transactions in the FX market. These uncertainties have diffusion form and are dealt with according to principles of portfolio optimization in continuous time. The price information content of the local order flow observed by the dealer, has to do with the existence of “global” investors in the dealer’s customer base. Global investors have access to bid-ask quotes of more than one dealer, along with superior information about fundamentals. These two factors help them generate better estimates of the future inter-dealer transactions than both the “local” investors (i.e. those who only trade with their single local dealer) and any of the dealers themselves. I obtain the formula for the dealer spread around the equilibrium price. I also show how informational customer-dealer asymmetries generate non-stationary cumulative order flows as well as deviations of the inter-dealer transaction price from uncovered parity of national asset returns.

JEL Classification: F31, G11, G29, D49, D82

1. Introduction

Closing the persistent gap between foreign exchange market microstructure models and traditional chapters of international macroeconomics is an ambitious theoretical challenge that has also become an increasingly important issue in the eyes of policy makers inside the central banks. The forex microstructure theory originally declared the intention to deal with the exchange rate formation in terms that could be recognized by financial market practitioners, as opposed to textbook macroeconomic wisdom that seems more often than not, to fail to be reflected accurately by FX-traders. The problem is especially urgent in the eyes of a monetary authority whose very *raison d’être* is the presumed ability to implement a desired macroeconomic objective through actions taken in the money and FX markets. At present, however, the best-known microstructure models operate with conceptual shortcuts and information-theoretic constructions that make their messages even more, not less, distant from the language spoken in the dealer rooms than standard propositions of classroom macroeconomics.³ A problem in itself is to derive from any of these models an empirically meaningful corollary testable on available data.

¹ Monetary Dept., the Czech National Bank, Na Prikope 28, CZ-115 03 Praha 1, and Inst. of Information Theory and Automation, Pod vodarenskou vezi 4, CZ-182 08 Praha 8, Czech Republic; e-mail: Alexis.Derviz@cnb.cz or aderviz@utia.cas.cz.

² Comments and valuable advice by Rich Lyons from U.C. Berkeley are gratefully acknowledged. The usual disclaimer applies.

³ The author’s experience is based on contacts with people from the open market operations departments of a number of reserve banks (his own, i.e. the Czech National Bank, the Bank of England, and a number of others

The present paper investigates the ability of a continuous time portfolio optimization model for FX dealers and investors, to highlight two groups of phenomena in the FX market. The first are the “long run” properties of the exchange rate, such as a (generalized) uncovered parity or competitive quoting by multiple dealers around a clearing price. The second group, which is more “short run” by nature, refers to information extraction procedures and Bayesian belief updating by participants in the inter-dealer market in the face of changing fundamentals.

In the course of our analysis of the short-long run factor interplay in the decentralized FX market, we shall focus on the “fundamental” property called the uncovered total return parity (UTRP) for the exchange rate. UTRP associates the exchange rate expectations for a given period with the difference in total returns (instantaneous dividend over price plus capital gain) on a pair of representative securities. These total returns coincide with yields to maturity in continuous time. Thus, returns quoted on a secondary market and not the money market rates which are pre-defined for the time interval in question, constitute the continuously updated measure of the expected move in the exchange rate (see Derviz, 1999b, for details). UTRP would be a natural result in, practically, any reasonably defined dynamic model of international asset pricing and shall replace the much-compromised uncovered interest rate parity of naïve no-arbitrage models. Empirical verification of UTRP also provides promising results. This parity theorem would be valid exactly and permanently in a Walrasian auctioneer setting which ignores microstructure, with a representative agent and markets clearing at each moment. However, when one studies the exchange rate formation in a dealership market, deviations from the fundamental UTRP come about as a natural consequence of agent heterogeneity. These deviations can be used to characterize the current state of information and inventory flows between dealers and investors. The proposed model establishes a link between an individually observed order flow, Bayesian filtering of imprecisely known fundamentals by the dealer, and a seemingly non-stationary behavior of the country premium in the UTRP formula. The result obtained in this paper is that deviations from UTRP are most pronounced at times when some dealers are learning new information from their clients through the observed order flow. During such periods,

1. the exchange rate and the inter-dealer order flow variability must be higher than under complete symmetric information
2. dynamic processing of new signals by the dealers leads to deviations from UTRP that persist until the new information is absorbed; under certain circumstances, Bayesian learning by the dealers can lead to a permanent revision of the country premium level in the uncovered parity relation.

In this sense, the model describes situations when the order flow effects are persistent.

In order to highlight the “right” price discovery process starting with the end-users of the foreign currency (investors) and going through the dealers who learn from them, the existing microstructure literature makes a sharp distinction between clients and dealers. This happens both in single dealer sequential trade models such as Glosten and Milgrom, 1985, or Easley

from emerging economies). The conclusion is, not surprisingly, that a dealer is usually well versed in the uncovered interest rate parity argument (and even knows that it does not hold). He/she can also easily comment on the signal impact of released fresh data about inflation or GDP, on the exchange rate movements. It is, however, totally unrealistic to expect a dealer to analyze the separating versus pooling nature of equilibrium in a trader-specialist game of Easley and O’Hara, 1987, or Glosten, 1989. Scholars may use the latter models to *analyze* the actions of real dealerships, but not to *communicate* with them.

and O'Hara, 1987, with a single risk-neutral market maker, and in multiple dealer models of the Evans and Lyons, 1999, type. This distinction is barely possible in practice, and most of available price and trade size data carry no stamp of the purpose of any given FX-transaction. More importantly, in real life, the roles of investor and dealer in the forex market are often combined in the performance of cooperating parts within one and the same firm. Most typically, a major commercial or investment bank has dealers who service its customers and participate in the inter-dealer market, but also operates proprietary desks that exercise FX trades for the needs of its own portfolio management. Access to electronic cross-border security trade and information systems is not a privilege of dealers either: the community of Reuters and/or Bloomberg quote screen users includes a considerable part of internationally active companies of very versatile profiles. Access to these facilities becomes highly attractive to any company whose cross-border operations attain a certain size.

In view of the given circumstances, empirically testable models of FX-trade cannot, and should not, rely on the strict client-dealer separation. The approach taken here assumes that access to the inter-dealer market is possible for a non-empty subset of investors, who are able to both search for the best dealer quotes and exercise a trade with the dealer they choose. The only distinction between such "global" investors and dealers in the proper sense is that the latter voluntarily take on the function of market makers and therewith - the commitment to provide two-way quotes and trade with anybody approaching them. Apart from that, dealers have the same intrinsic motive for holding foreign currency positions as any other investor, since they are active in the same lines of international business as the latter. It is this synthetic "dealer-investor" that we are going to model.

The majority of forex microstructure models dealing with asymmetric information define the signals as unbiased estimates of some "true" or "fundamental" value of the exchange rate. This is a feature they have inherited from microstructure studies of other financial market segments (e.g. equities, where this true value is associated with revealed earnings or dividend data). However, in the case of the exchange rates, no such revelation of absolute truth exists. Therefore, a careful choice of the content of the signals, as well as a justification of its relevance for the price setting, is needed. In the present model, this task is fulfilled by letting the signals reaching the dealers contain information about the "other side of the market", i.e. the aggregate behavior of the receiving side of their outgoing trades. This allows us to close the model in a natural way and derive some useful properties of the arising equilibria.

Another feature of existing models of inter-dealer market that complicates empirical applications, is the strict sequencing of information-laden events in them. (Cf. Evans and Lyons, 1999: first a round of dealer trades with customers signaling fundamental information, then an inter-dealer round to share inventory risk evenly between them, then another round of customer trades to unwind open positions.) On the contrary, data on FX trades of a given financial institution reflect a random sequence of events pertaining to own investment needs, the inter-dealer and the customer-dealer operations. In other words, the three named categories of transactions generate three simultaneously evolving processes. At each moment, the dealer must exercise a belief update based on what he has observed so far. He is hardly able to say which trading round informs him about what in particular. Under such conditions, a continuous time stochastic model of the dealer behavior appears to be the most convenient.

A non-negligible argument in favor of a continuous time dealer model with diffusion uncertainty is its better analytical tractability compared to discrete time models with arbitrary statistics of risk factors. There already exists a line of literature that develops the findings of

the Kyle, 1985, risk-neutral market maker model. For example, Back, 1992, Back and Pedersen, 1998, and Back, Cao and Willard, 2000, work with price and cumulative orders in a security market (applicable to forex) in semimartingale form, with diffusion components originating in the action of noise traders. When one goes over to optimizing models for risk-sensitive agents, the stochastic maximum principle becomes an even more powerful tool of analyzing equilibrium price and order flow dynamics in continuous time, once the dealer optimization problem is properly defined in this setting. The crucial challenge here is to identify and interpret the information asymmetry and Bayesian belief updating phenomena and their role in the obtained solutions.⁴ This task in a continuous time portfolio model of the Cox-Ingersoll-Ross type with heterogeneous beliefs was first addressed by Detemple, 1986. Zapatero, 1998, used the same approach to derive results on asset price volatility under asymmetric information in continuous time.

Since each dealer enjoys a local customer base that trade with him exclusively, there is enough space for a non-trivial bid-ask spread in his quotes. The well-known reason for existence of spreads, identified in most models with informational heterogeneity of traders, namely the compensation of losses from trades with the informed, by profits made in trades with the uninformed, is implicitly present in this setting as well. Obviously, the role of the informed traders is played by the global, while that of the uninformed – by the local traders. The latter, when they trade at given spreads, are subject to the exercise of monopoly power by their dealer. This understanding is close to the ideas of Copeland and Galai, 1983, and Perraudin and Vitale, 1996. However, the present model is able to come up with an explanation of non-trivial – and variable – bid-ask spreads even when there is more than one dealer in the market. The existence of global investors extends the competition from pure inter-dealer interaction to the sphere of dealer-customer relations. In this respect, they seem to be a key factor in the function of a decentralized inter-dealer market, which is typical for forex trades.

The existence of informational heterogeneity between investor groups of different residences is pivotal in generating non-trivial order flows and additional sources of exchange rate volatility in our model. That is why it is important to note that such heterogeneity has already found reflection in finance literature. For instance, Brennan and Cao, 1997, argue that asymmetry between domestic and foreign investors in the knowledge of domestic asset returns has an impact on the direction and volume of cross-border equity trade. More specifically, a recent working paper by Seasholes, 2000, argues that large foreign investors in emerging markets have an informational advantage and cash in higher returns on domestic equity, over the majority of domestic investors. (Therefore, cross-border informational asymmetry does not always mean an advantage for the residents.) Our model shares the Brennan and Cao thesis that residency-based informational heterogeneity matters, but also allows for difference in informational endowments between local and global investors, in accordance with the Seasholes conjecture.

The paper is structured as follows. Section 2 describes the model variables, the structure of uncertainty and the dealer's optimization problem (Subsection 2.1), after which the optimal dealer strategies for passive (at own quotes) and active (at other dealers' quotes) FX trades are derived (Subsection 2.2). I also show how a non-zero bid-ask spread and a competitive mid-

⁴ The previously mentioned continuous time models of FX trade by Back and colleagues are based on Kyle's auctioneer and do not include explicit dealers. Besides, these models abstract from the usual Brownian asset return uncertainties faced by an investor and concentrate on the market interpretation of narrowly defined exogenous signals about the foreign currency value.

quote arise as a result of dealer optimization. Section 3 analyzes the long-run FX transaction price dynamic resulting from the uncovered total return parity. Section 4 discusses order-flow interpretation and updating of beliefs by dealers, leading to deviations from uncovered parity of returns. Section 5 concludes and indicates possibilities of future developments of the model.

2. The investor-dealer decision problem

2.1 The model

The world economy is split into the home country with legal tender M and the foreign country with legal tender I . By the exchange rate we understand the price of I in M -terms. Investors resident in the home country account their operations in M , those resident in the foreign country use the accounting unit I .⁵ Among the investors (of both residencies) there is a subset of those who offer forex dealer services to the others. That is, they agree to provide an ask quote p^a which is the price at which they agree to sell I in exchange for M , and a bid price p^b at which they agree to buy I in exchange for M . We will call them investor-dealers or simply dealers. Each dealer has a local customer base of those investors who only exercise forex trades through him at given quotes (they are free to vary the traded quantity from zero to infinity depending on the quotes they see). Beside them, there are other (global) investors who are able to search for the best quote among all dealers, at a cost. Every dealer himself has the ability of costly discovery of the best quote among other dealers. One of the distinctions between investor-dealers and investors without the dealing capacity is that the former category faces a lower cost of quote search. Another one is the priority of the non-dealer investors in learning the news about fundamentals, to be discussed in Section 4. In the case of a smaller FX market (e.g. a European currency not belonging to the euro zone or an emerging economy currency) one can loosely associate global investors with multinationals and international financial corporations present in the local economy.

Altogether, the order flow that a dealer observes consists of three parts. The first is the supply of I or demand for I from his local customers, dependent only on his own quote values p^a and p^b and the economic fundamentals influencing the customer base. This flow is unobserved by other parties. The remaining two order flow components come from the inter-dealer market. These are the orders of those (a) global investors and (b) other dealers, whose quote search has resulted in the choice of this dealer for the desired trade in the given period. For every party involved, participation in the inter-dealer market is the source of an information update about both the global fundamental variables and the future exchange rate moves, read off the received trades. The inter-dealer market is partially decentralized in the sense that each participant only observes his/her own trades, whereas the knowledge of standing best quotes is the result of a costly search for each outgoing trade. These search costs will be incorporated into the overall transaction costs (to be defined in a separate paragraph below). One can think of an agent with access to electronic and voice brokers who service parts of the market, but who is unsure whether a better price cannot be found by approaching certain dealers individually. While not modeling these tradeoffs between broker and individual search explicitly, we adopt the understanding that the best price can always be found, but the resources dedicated to the search grow faster than the desired transaction amount.

⁵ The model to be defined will accommodate the Jensen inequality-related properties of the expected exchange rate changes (the residency-based asymmetry of the latter in a naive no-arbitrage setting is known as Siegel's paradox). Thus, different residencies will not cause formal contradiction. See Derviz, 1999b, for the details of Siegel's paradox resolution in the corresponding class of models.

The international investor exercising the function of an FX-dealer will be modeled as an agent characterized by four state variables. These are: x^0 – domestic cash holdings, x^d – holdings of the composite domestic asset, x^i – foreign cash holdings, x^f – holdings of the composite foreign asset. The domestic asset pays out a random rate of return dr^d in M-units. I also assume that maintaining the investment portfolio x^d requires continuous inputs of funds (management costs) without which the value x^d deteriorates at a stochastic rate $d\mathbf{a}^d$. (This will be convenient since one can then consider equilibria with constant average levels of asset holdings.) Let one unit of x^d cost X^d units of M. Analogous values for the foreign investment portfolio are denoted by dr^f , $d\mathbf{a}^f$ and X^f , respectively (the latter price is, naturally, in I-units).

It is convenient to assume that the cash holdings x^0 and x^i earn their own rates of return dr^0 and dr^i , respectively. In the simplest variant, it can be the overnight money market rate (or its proportional part if the elementary period is intra-day). More realistically, one should include in the cash variables the inter-bank loan/deposit positions and FX swap positions. In that case, dr^0 (dr^i) is the instantaneous rate paid in the domestic (foreign) money market on the corresponding portion of x^0 (x^i), i.e. the generator short rate of the domestic (foreign) term structure.

For any strictly positive Ito process z , symbol \widehat{dz} will be a shorthand for dz/z . The drift and diffusion coefficients of the asset return and price processes introduced above will be denoted as follows:

$$\begin{aligned} dr^0 &= i^d dt + I^d dZ, \quad dr^i = i^f dt + I^f dZ, \quad dr^d = n^d dt + N^d dZ, \quad dr^f = n^f dt + N^f dZ, \\ d\mathbf{a}^d &= a^d dt + A^d dZ, \quad d\mathbf{a}^f = a^f dt + A^f dZ, \\ \widehat{dX}^d &= \mathbf{p}^d dt + \mathbf{r}^d dZ, \quad \widehat{dX}^f = \mathbf{p}^f dt + \mathbf{r}^f dZ. \end{aligned}$$

Since it is not the purpose of this paper to study GARCH or other non-constant variance effects in asset prices, all diffusion coefficients I, N, A and \mathbf{r} are assumed constant.

Ito processes $r^0, r^d, r^i, r^f, \mathbf{a}^d$ and \mathbf{a}^f belong to the exogenous sources of uncertainty in the model. Jointly, these processes, together with two more to be introduced in Subsection 2 and Section 4, respectively, generate the information filter $\mathbf{F}=(F_t)_{t \geq 0}$ to which all stochastic processes appearing in the model will be adapted. (F_t denotes the partition of the event space corresponding to the full information about the operation of the markets, on which the objective probability measure valid at time t is defined.) Let the diffusions adapted to \mathbf{F} be spanned by a vector Z of mutually independent standard Brownian motions.

The current FX transaction price As mentioned above, it is assumed that both the global investors and the investor-dealers are allowed to search in the inter-dealer market for acceptable quotes. Accordingly, let us introduce the notion of the best attainable bid and ask prices for a given dealer, P^b and P^a . These are good for one unit of I. Trade size-dependent non-linearities will be accommodated by means of convex transaction costs (see below).

Assumption 1 For each global investor and each dealer, the best individually attainable (highest) bid price is equal to the best attainable (lowest) ask price: $P^b = P^a = P$.

The difference between the best/standing ask and bid is sometimes called “inside spread”, or “touch”. Assumption 1 sets this inside spread to zero. The rationale of this assumption lies in the natural information-dissemination ability of the inter-dealer market and competitiveness of its participants. Indeed, if one observed $P^b > P^a$, then there would be a clear arbitrage opportunity. In the opposite case $P^b < P^a$ there would be no inter-dealer trades at all until the

too high asks went down and/or the too low bids went up, to meet each other and provide gains from trade to those who seize the opportunity first. That is, we believe that there will always be an agent who will see the deadlocked market and undercut/overbid the standing quote in his favor. As soon as all the arbitrage and market share appropriation possibilities have been exhausted, there remains a group of inter-dealer market sellers and another group of buyers. They transact at the mutually acceptable price P . The buyers' ask price and the sellers' bid price quotes are non-competitive: above P in the first group and below P in the second. Therefore, the buyers do not have to sell and the sellers do not have to buy.

To back up the legitimacy of Assumption 1, first note that the inside spread diminishing with the number of competing dealers is one of the outcomes of the model by Ho and Stoll, 1983. Their model also predicts that spreads of individual dealers would not fall to zero even when their number becomes large. This is in line with the results to be obtained in the present paper. On the empirical side, Goodhart et al., 1996, observe that interdealer FX spreads on major currency pairs are usually tiny (as small as 1 pip or 1/10 of the basis point for standard amounts). However, these are the numbers reported by the Reuters D2000-2 system of electronic brokerage, i.e., again, the inside spreads, giving a perfect match with Assumption 1.

The dealer population is split into sellers and buyers at every given moment. The most trivial reason why someone becomes a seller and someone else – a buyer is the heterogeneity of asset endowments. But, even in the absence of the latter, there is a possibility of differences in privately accounted net returns on assets between dealers. For instance, foreign residents may have a handicap with regard to attrition in their domestic asset holdings, and vice versa. Domestic dealers may enjoy a privileged access to the domestic money market, with a resulting higher return on domestic cash balances, and a symmetric advantage may be enjoyed by foreign residents with respect to foreign cash. Other types of heterogeneity can be imagined as well, including those that vary over time. This paper will focus on heterogeneity of beliefs.

The last-observed inter-dealer transaction price P is the one used in the dealer's accounting. Each agent views it as a strictly positive Ito process with coefficients defined by

$$\widehat{dP} = \mathbf{p}dt + \mathbf{r}dZ \quad .$$

Investment performance measure For the later purpose of defining preference over the paths of investor-dealer's actions, we are introducing the notional dividend rate \mathbf{d} to be accounted for in every infinitesimal period dt . Our model shall be able to work with very short time steps (such as those of the real-time data of electronic FX-trade systems). In this case, the defined dividend rate shall be understood as an imaginary infinitesimal contribution of the dealer department to the dividend fund of the firm, whose integral over longer periods, in reality, is being withdrawn from the cash balance at discrete intervals. Having this infinitesimal dividend rate in the model is a convenient way to make the dealer accountable for the ultimate performance of the FOREX activity of the firm, and, at the same time, bring the preference structure of the agent close to the standard portfolio optimization setups.

Price search costs In order to generate non-trivial supply and demand schedules (for instance, Glosten, 1989, uses non-linear pricing schedules in his investor-dealer game in both monopolistic and competitive dealer settings) facing the other side of the market, it is usual to define transaction costs that grow in a convex way in the transaction volume. It is well known that observable transaction costs incurred by a big real-life investor, let alone a dealer, are

usually negligible for small/standard amounts, but grow rapidly when the trade size exceeds the standard. Accordingly, non-linear pricing schedules are a reality. Beside that, some form of generalized transaction costs seems to lie behind the dealer behavior, even if they are not readily measured. In this paper, we take the view of these costs arising from the quote search in the inter-dealer market. Namely, it is logical to assume that in order to trade volume $v^i dt$ of I at price P (which is the last observed highest bid if $v^i < 0$ and the last observed lowest ask if $v^i > 0$), one must pay $Pk(v^i)dt$ units of M . (This sum becomes a receipt and enters the balance with a negative sign, if I is sold, i.e. $v^i < 0$.) Here, k is a strictly increasing and strictly convex function with $k'(0) = 1$. The most natural example would be a quadratic function of the form

$$k(v) = v + k_0 v^2 / 2.$$

For reasons of consistency, one should assume the same search cost mechanism to be present in the two other market segments as well. That is, if amounts v^d and v^f of the domestic and the foreign asset are purchased per period (sold if negative), the M - and I -balances are reduced by amount $X^d k(v^d)dt$ and $X^f k(v^f)dt$, respectively.

The dealer's order flow As was mentioned in the introduction, it is extremely difficult to obtain separate data on the customer and the inter-dealer components of a dealer's order book. Instead, we distinguish between the local customers of the given dealer (do not have access to or exercise FX trades with others) and the participants of the inter-dealer market. Roughly, the latter set includes all those who access the same screen. Both the investors who search among available dealer quotes ("global investors"), and dealers themselves, belong there. Accordingly, the pair of ask and bid I -prices p^a, p^b , announced by the dealer, shall service everyone at the moment of announcement. However, we assume that the dealer is able to distinguish between a locally riskless (non-diffusion) demand $D(p^a)dt$, a locally riskless supply $S(p^b)dt$, and a random flow $x^0 dL/P = x^0 (ldt + IdZ)/P$, which can go both ways, in his order flow. Altogether, the accounted change in the dealer's M -balance, induced by trades at his quotes, is equal to

$$\left(p^a D(p^a) - p^b S(p^b) \right) dt + x^0 (ldt + IdZ) = \left(p^a D(p^a) - p^b S(p^b) \right) dt + x^0 d\Lambda, \quad (1)$$

while the corresponding change in the I -balance is

$$\left[S(p^b) - D(p^a) \right] dt - \frac{x^0}{P} (ldt + IdZ) = \left[S(p^b) - D(p^a) \right] dt - \frac{x^0}{P} d\Lambda. \quad (2)$$

Observe, in particular, the last diffusion terms in (1) and (2): $x^0 IdZ$ and $-\frac{x^0}{P} IdZ$. They reflect the dealer's uncertainty about the order flow from both the global investors and other dealers. Indeed, one can generate the dealer's diffusion-type order flow as a limit of two small-interval discrete flows of very small purchases at p^a and very small sales at p^b with a non-zero degree of randomness in the direction of the trade and its exact magnitude. Then the drift parts of the limit flows will correspond to the demand and supply that are certain during the infinitesimal interval dt , namely $\left[D(p^a) + \frac{x^0 l^a}{P} \right] dt$ and $\left[S(p^b) - \frac{x^0 l^b}{P} \right] dt$, where $l^a + l^b = l$.

The diffusion part reflects the remaining uncertainty. Considering the dealer an \mathbf{M} -resident, we account for the random order flow in \mathbf{M} -terms by expressing it as a stochastic growth rate of the \mathbf{M} -cash account x^0 . It can be shown that the \mathbf{I} -position in that case evolves exactly as is shown by the last term in equation (2)⁶. The presence of x^0 -proportional part $-x^0 \frac{l}{P}$ in the drift component of the order flow is explained by the possibility of the probability revision by the dealer in accordance with the Girsanov theorem. The consequences of such a probability measure change in the case of a Bayesian update of beliefs will be discussed in Section 4.

If the dealer were a foreign resident, the uncertain parts of the order flow that she faces would be x^i -proportional, and the certain FX supply- and demand-entries in equations (1), (2), coming from the local customers, should be modified accordingly.

Summing up the definitions given above, we come to the following system of state-transition equations for the state vector $x=[x^0, x^d, x^i, x^f]^\top$ of an international investor with the dealer function (with domestic residence):

$$dx^0 = x^0 dr^0 + x^d dr^d - \mathbf{d}dt - X^d k(v^d)dt - Pk(v^i)dt + (p^a D(p^a) - p^b S(p^b))dt + x^0(l dt + \mathbf{I} dZ), \quad (3a)$$

$$dx^d = -x^d d\mathbf{a}^d + v^d dt, \quad (3b)$$

$$dx^i = x^i dr^i + x^f dr^f - X^f k(v^f)dt + v^i dt + [S(p^b) - D(p^a)]dt - \frac{x^0}{P}(l dt + \mathbf{I} dZ), \quad (3c)$$

$$dx^f = -x^f d\mathbf{a}^f + v^f dt. \quad (3d)$$

Preferences The model in continuous time that is being defined in this paper, is intended to be applicable to high frequency, e.g. intra-day, data. With such short time steps, it is not obvious whether a time preference/discount factor of the standard optimization models is applicable. We have already agreed that the dividend rate \mathbf{d} in equation (3a) is to be understood as an instantaneous contribution to the dividend fund, to be integrated over a finite time period. Let there exist a function u satisfying the conventional growth and concavity conditions, such that $u(\mathbf{d})$ measures the period utility (of the shareholders), derived from receiving contribution \mathbf{d} to the dividend fund. Similarly, infinitesimal rates dr and $d\mathbf{a}$ of asset return and deterioration in the model can be understood as sums of the per-period contributions to the integral expected returns over a finite period (the drift parts) and unexpected innovations in these integral returns (the diffusion parts). Accordingly, an instantaneous time preference rate can be defined as an infinitesimal contribution to an integral discount factor valid for a finitely distant period in the future. Another possibility, to be exploited here, is to understand this rate, denoted by \mathbf{b} , as the parameter of a Poisson “death” process. Since the present model neither requires a dealer to close the FX position in a predetermined finite time, nor accepts the existence of an exogenous final/underlying value of the currency, the following construction has the objective of replacing the artefact of the “true liquidation value under discovery”, often utilized in microstructure models.

⁶ The said transition to the limit of discrete processes is a generalization of the well-known procedure of generating a geometric Brownian motion as a limit of random walks with a drift, when the step size goes to zero, see e.g. Dixit and Pindyck, 1994, Section 3.2.

For each time moment t , the dealer will have to close down his forex trade business within the next time interval dt , with probability $1-e^{-bdt}$, and liquidate the outstanding engagement in currency I at current prices. The result of liquidation, i.e. the balance $x^0 + Px^1$, is evaluated by means of a strictly increasing and concave *exit (or bequest) utility* function G . As a result, at any time t the domestically resident dealer maximizes

$$E \left[\int_t^\infty e^{-b(s-t)} \{u(\mathbf{d}_s) + \mathbf{b}G(x_s^0 + P_s x_s^1)\} ds | F_t \right] \quad (4)$$

with respect to control path $s \mapsto (\mathbf{d}_s, v_s^d, v_s^i, v_s^f)$, $s \geq t$, subject to (3), given the current values x_t of asset holdings. Note the appearance of parameter \mathbf{b} in front of G in the integrand in (4). Since, in each period ds , the firm liquidates with probability $1-e^{-bds}$, the expected utility derived from the liquidated position is equal to

$$(1 - e^{-bds}) G_s = \frac{1 - e^{-bds}}{ds} G_s ds \approx \mathbf{b} G_s ds.$$

If the dealer resides abroad, her liquidation balance entering G shall be accounted for in I-units and be defined as $\frac{x^0}{P} + x^i$.

2.2 Optimal policies of the dealer

Following the results of the Appendix, we can characterize the solution of the problem (3), (4) by means of the shadow prices \mathbf{x} of the four held assets, which are the adjoint processes of the problem appearing in the maximum principle. Given the Hamiltonian of the problem, as calculated in the Appendix, the optimal actions of the dealer are characterized by the following first order conditions:

$$\mathbf{d} = g(\mathbf{x}_0), \quad (5a)$$

$$k'(v^i) = \frac{\mathbf{x}_i}{\mathbf{x}_0 P}, \quad (5b)$$

$$k'(v^d) = \frac{\mathbf{x}_d}{\mathbf{x}_0 X^d}, \quad (5c)$$

$$k'(v^f) = \frac{\mathbf{x}_f}{\mathbf{x}_i X^f}, \quad (5d)$$

$$p^a \left(1 - \frac{1}{\mathbf{e}^a} \right) = \frac{\mathbf{x}_i}{\mathbf{x}_0}, \quad p^b \left(1 + \frac{1}{\mathbf{e}^b} \right) = \frac{\mathbf{x}_i}{\mathbf{x}_0}, \quad (5e)$$

where g denotes the inverse function to u' , $\mathbf{e}^a = \frac{-p^a D'(p^a)}{D(p^a)}$ is the price elasticity of the

local I-demand and $\mathbf{e}^b = \frac{p^b S'(p^b)}{S(p^b)}$ is the price elasticity of the local I-supply, both observed by the dealer.

Further, according to the adjoint equation stated in the maximum principle of the Appendix, the shadow asset prices faced by the dealer satisfy the system of equations

$$d\mathbf{x}_0 = \mathbf{x}_0 \left(\mathbf{b}dt - dr^0 - ldt - \mathbf{I}dZ + \left| \mathbf{I}^d + \mathbf{I} \right|^2 dt \right) + \mathbf{x}_i \left(\frac{ldt + \mathbf{I}dZ}{P} - \frac{(\mathbf{I}^d + \mathbf{I} + \mathbf{I}^f) \cdot \mathbf{I}}{P} dt \right) - \mathbf{b}_G' dt, \quad (6a)$$

$$d\mathbf{x}_d = \mathbf{x}_0 \left(-dr^d + (\mathbf{I}^d + \mathbf{I} - A^d) \cdot N^d dt \right) + \mathbf{x}_d \left(\mathbf{b}dt + d\mathbf{a}^d + \left| A^d \right|^2 dt \right) - \mathbf{x}_i \frac{N^d \cdot \mathbf{I}}{P} dt, \quad (6b)$$

$$d\mathbf{x}_i = \mathbf{x}_i \left(\mathbf{b}dt - dr^i + \left| \mathbf{I}^f \right|^2 dt \right) - \mathbf{b}_G' dt, \quad (6c)$$

$$d\mathbf{x}_f = \mathbf{x}_i \left(-dr^f + (\mathbf{I}^f - A^f) \cdot N^f dt \right) + \mathbf{x}_f \left(\mathbf{b}dt + d\mathbf{a}^f + \left| A^f \right|^2 dt \right). \quad (6d)$$

Naturally, if a dealer resides in the foreign country, then she faces the shadow prices whose laws of motion must be a symmetrically adjusted version of (6), in accordance with her units of account. In other words, in equations (6a)-(6d), one would have to switch the index pairs (0,d) and (i,f). Also, the term analogous to $\mathbf{I}dZ$ and other terms containing \mathbf{I} would appear in equations (5c), (5d) instead of (5a), (5b). Only equation (5e), characterizing the quote setting, would remain as it is.

Spread An immediate consequence of the f.o.c. (5e) is the following expression for the dealer's bid -ask spread:

$$\frac{p^a}{p^b} = \frac{1 + \frac{1}{\mathbf{e}^a}}{1 - \frac{1}{\mathbf{e}^b}}. \quad (7)$$

If elasticities \mathbf{e}^a and \mathbf{e}^b are high, the log of the right hand side of (7) is approximately equal to $\frac{1}{\mathbf{e}^a} + \frac{1}{\mathbf{e}^b}$. The conjecture about high price elasticities of demand and supply on the side of the dealer's customer base seems justified, since the customers are not facing a unique dealer-monopolist. It is reasonable to assume that as soon as the disadvantage of trading with a monopolist or monopsonist becomes too evident, a customer can always try to look up another one, i.e. turn "global" despite the associated costs. In the sequel, I will assume the existence of a common upper bound for inverse elasticities $\frac{1}{\mathbf{e}^a}$ and $\frac{1}{\mathbf{e}^b}$ for all customer bases.

Equation (7) would come up in a standard way in any monopolist two-way market maker problem, regardless of the presence of uncertainty. However, the model defined here contains two less standard elements. First, it is able to generate spreads that are variable in time and stochastic. Second, by considering a dealer who participates in a competitive inter-dealer market, the model restricts the equilibrium paths of dealer quotes to those that reflect information about fundamentals, to the extent the latter is disseminated by other parties' actions (see Section 4 below).

To account for the observed variable FX spreads, it is necessary to consider only such supply and demand schedules faced by the dealer, that generate non-constant non-deterministic price

elasticities. Typical functional forms for functions S and D , indeed, satisfy this requirement. To see this, observe that any investor can be modeled by means of equations analogous to (3) and (4) above, if one eliminates the dealer order flow parts from transition equations (3a) and (3c). Then the FX supply/demand schedule will be given by an equation formally identical to (5b). Depending on the current values of shadow prices relative to the offered market I-price, some investors will be sellers and other – buyers. In both cases, the aggregate S and D that the dealer sees will be linear functions of the inverse price, with stochastic coefficients. Such functions give rise to variable stochastic price elasticities.

The laws of motion (6) of the shadow asset prices characterize them in terms of the *fundamental variables* of the defined economy. These fundamentals in the present model are comprised of various components of the total asset returns (dividends, price movements, depreciation rates), plus a variable which characterizes the aggregate flow of I-funds to/from the domestic dealers (to be defined in Section 4). The fundamental information accessible to different groups of market participants can have different quality. Possibilities to model information dissemination processes with regard to fundamentals within the present approach will be discussed in Section 4. But first, here, to see that dealer quotes must reflect the processed information on fundamentals competitively (therewith reducing monopolistic welfare loss effects), it is necessary to deduce a number of properties of possible market equilibria. This is done in the next section.

3. Shadow price parity, uncovered parity of total returns, and the asymptotic dynamic of the exchange rate

The first order conditions of optimality (5) can be used to state a fairly general property of the exchange rate dynamics, which can be called the generalized uncovered total return parity condition. The underlying property of equilibrium in the asset markets is the same as the one that leads to the consumption-based CAPM. A similar result would obtain in most international portfolio optimization models. In contrast to the very much discredited uncovered interest rate parity of the textbooks, the total return parity enjoys enough empirical support (see Derviz, 1999b, for a model and empirical verification; another approach with more data on leading currencies can be found in Nada-De Simone and Razzak, 1999).

To formulate the result, it is necessary to define the total instantaneous rates of return of the domestic and the foreign asset:

$$dR^d = \frac{dr^d}{X^d} + \widehat{dX}^d - d\mathbf{a}^d = y^d dt + \mathbf{u}^d dZ, \quad dR^f = \frac{dr^f}{X^f} + \widehat{dX}^f - d\mathbf{a}^f = y^f dt + \mathbf{u}^f dZ.$$

Next, define the auxiliary variable Y as

$$Y = \frac{\mathbf{x}_f X^d}{\mathbf{x}_d P X^f}.$$

Variable Y can be called the *shadow price parity* index. The reason is the fact that, according to Ito's lemma applied to equations (6b) and (6d) for the shadow prices, its stochastic differential satisfies the following property:

$$\widehat{dY} = dR^d - dR^f - \widehat{dP} + Adt. \quad (8)$$

Here, A (the *disparity* term) is the sum of a number of covariance terms (they come about as a consequence of Ito's lemma). They are usually small and, according to our assumption on the constancy of diffusion coefficients of exogenous variables, A is a constant.

The first three terms on the right hand side of (8) define the *uncovered parity of total returns* (UTRP) for the exchange rate P . Namely, if Y were a constant and the disparity term A —close to zero, then the relative expected change in P between times t and $t+dt$ would be equal to the total return differential. Thus, the uncovered return parity can be formulated as the equality of the exchange rate differential to the instantaneous return rate differential between representative domestic and foreign securities plus a covariance term:

$$\widehat{d}P = dR^d - dR^f + A dt. \quad (9)$$

A priori, there is no reason to see the shadow price parity index as a constant. Instead, by introducing the shortening notations $k^d = k'(v^d)$, $k^i = k'(v^i)$, $k^f = k'(v^f)$, we can derive from (5b)-(5e) the following expression for Y :

$$Y = \frac{k^i k^f}{k^d}. \quad (10)$$

If the markets were characterized by the existence of a representative agent with no dealer-customer distinction (and the clearing prices were set by a Walrasian auctioneer), then all security purchase rates would have to be set to constants in equilibrium. Possible equilibria would then include ones with constant asset price trends, constant Y and the exchange rate that satisfies the uncovered return parity (9) exactly.

To analyze equilibria in the presence of dealers, it is useful to make a simplifying assumption about the behavior of the asset market segments other than the forex. Specifically, assume that the asset holdings x^d and x^f of every dealer and investor possess long run average limit levels \bar{x}^d , \bar{x}^f . Each agent simply maintains the long run average level of both asset holdings in the portfolio by compensating, in the drift part, for their continuous attrition described by (3b) and (3d). Then the optimal purchase rates v^d and v^f are positive constants:

$$v^d = \bar{v}^d = \bar{x}^d a^d, \quad v^f = \bar{v}^f = \bar{x}^f a^f,$$

making the asset holding processes revert to \bar{x}^d , \bar{x}^f in the mean:

$$dx^d = a^d (\bar{x}^d - x^d) dt + x^d A^d dZ, \quad dx^f = a^f (\bar{x}^f - x^f) dt + x^f A^f dZ.$$

Regarding the price processes X^d and X^f , we shall assume that they are pure diffusions ($\mathbf{p}^d = \mathbf{p}^f = 0$).

The other three assumptions are aimed at limiting the long run price trajectories in the forex to the class of bounded ones. That is,

- a. only equilibria with the exchange rate growth rate \mathbf{p} bounded from both sides, will be considered;

- b. the attention is restricted to economies with identical preferences and processing cost functions of the dealers-domestic residents, dealers-foreign residents, global investors and local customer bases of each dealer (identity in each group);
- c. sizes of the client currency supply and demand sides, generated by non-dealer investors are bounded.

The benefit of the above restrictions is the existence of common equilibrium upper and lower bounds for individual inter-dealer orders v^i and, consequently, individual parameters k^i as well. Now the optimal quote equations (5e) offer a result about the long run behavior of the dealer quotes.

Proposition *Under the assumptions a.-c. made above, both bid and ask quotes are asymptotically equivalent to*

$$\left(1 \pm \frac{1}{e^{b,a}}\right)^{-1} Y^P . \quad (11)$$

In the above formula, the shadow price parity index Y is asymptotically a constant.

The necessity for Y to be almost constant is clear from the existence of upper and lower bounds for k^i . The expression (11) itself results from substituting into (5e) the following expression for the ratio $\mathbf{x}_i/\mathbf{x}_0$:

$$\frac{\mathbf{x}_i}{\mathbf{x}_0} = Y \frac{k^d}{k^f} P , \quad (12)$$

which follows immediately from (5b)-(5d).

The message of the above proposition is the existence of a strong link between the dealer behavior towards his/her clients and the constraints coming from the inter-dealer market. Although it is optimal for the dealer to maintain a positive spread between bid and ask quotes, the level on which the mid-point between bid and ask is set, is pinned down by the ratio of the dealer cash shadow prices. The latter are tied to the shadow prices of domestic and foreign assets or, more specifically, to the product of the last observed FX-clearing price and the shadow price parity index. In other words, the shadow price parity index governs the “quote shading” in the model. This index is fundamental (cross-border asset return differential-) driven and is, therefore, competitively determined by all dealers and investors. Accordingly, no dealer has an absolute monopoly power over the clients.

4. Fundamental information extraction from the order flow

As is usual in client-dealer models of FX trade, dealers in the present setup are supposed to learn from publicly observed events and private signals contained in the order flow. In the present model, Bayesian character of belief updating is implied in the dealer optimization procedure itself, as follows from the Girsanov theorem applied to the probability change in optimal control of diffusions (see Elliott, 1982, for details). However, it is impossible to predict unambiguously the direction of the exchange rate revision perceived from the outside, as a result of a particular news arrival and the corresponding belief update. Only the exact specification of the filtering procedure by the dealer can give a satisfactory answer. In this

section, we obtain a more specific answer about the equilibrium transaction price by making a more specific assumption about the nature of beliefs and signals. As a result, one will be able to formulate narrower results about the impact of belief changes regarding the asset return and the aggregate I-funds flow, on the dealer-investor actions.

In general, any Bayesian belief update in the present model must be characterized by a change in the probability measure Pr that a dealer uses in the optimization problem (3), (4). Under the new probability Pr^* , process Z spanning the exogenous uncertainties of the economy is no longer Brownian. There is another vector, Z^* , of standard mutually independent Brownian motions under Pr^* , related to Z in accordance with the Girsanov theorem. In its “naive” form, it reads

$$dZ = hdt + dZ^*, \quad (13)$$

where h is an \mathbf{F} -adapted continuous process satisfying a number of regularity conditions. The equilibria of the model can be associated with individual trajectories of h that generate the inter-dealer market order flow trajectories v^i , directed by purchasing dealers towards selling ones (see the dealer’s decision problem in Subsection 2.1). Every h also generates the trajectories of the dealers’ bid and ask quotes in the buying and the selling groups, and, thereby, the client order flows.

The complete statement of formation mechanism of h requires a specification of the exact filtering technique that the agent uses when processing the observations. The most popular would be, naturally, the Kalman-Bucy filter. I shall start by stating the basic properties that require a minimal amount of algebra but illustrate the dealer’s way of working with the order flow. It will have to do with the updating of beliefs on the exogenous variables’ drift terms after a new information about the current state of the economy has been registered and reflected in the customer order flow.

First, one shall posit the rule of client signal interpretation by the dealer. To do this, I introduce an additional source of exogenous uncertainty in the model, the random process M of the cumulative net outgoing purchases of foreign currency in the inter-dealer market by dealers – domestic residents. In other words, M is a measure of the cumulative inter-dealer order flow from M-residents to I-residents. The drift and diffusion coefficients of M are defined by

$$dM = mdt + \mathbf{m}dZ.$$

Given the homogeneity of the national dealer populations (Assumption b of Section 3), we conclude that m is formed by summation of identical active inter-dealer trades v^i across the M-resident dealers.

Let us define the *observations process* for the economy as the vector with components – individual sources of subjectively perceived uncertainty in the model. That is, define the vector diffusion process Q by $Q = [r^d, X^d, \mathbf{a}^d, r^f, r^f, X^f, \mathbf{a}^f, r^i, M]^T$. Symbolically, the evolution of Q will be written as $dQ = qdt + QdZ$. This is the dynamics under the objective probability Pr , utilized by those who are able to identify the drift part qdt precisely. (To be able to identify Q , it is sufficient to know only the *values* of Q across time, since, with this knowledge, the quadratic variation of Q can be computed. The following assumption makes use of the fact.)

Assumption 2 Every dealer observes the current state of the system, as expressed by Q , correctly, and all dealers agree upon the values of diffusion coefficients Q . However, dealers have imprecise information about the value of the drift term q . Consequently, an individual dealer's information filter F^Q generated by Q (and completed to satisfy the usual conditions) is cruder than the complete information filter F introduced in Section 2. Also, the dealer's subjective probability measure Pr^* differs from Pr .

As regards the first eight components of Q (I shall denote them by Q'), the assumption of their observability by any dealer is plausible enough and does not require extended justification. On the other hand, it might seem unnatural for a single domestic/foreign dealer to know exactly how much of I-currency his/her compatriot population has accumulated in aggregate. Assumption 2 is, in fact, weaker. It only requires that a noisy signal comprising (a) the full list of diffusions spanning the diffusion part of M , and (b) an imprecise measure of its drift, is received at every moment. This is made possible by observations of the order flow, as explained below.

Let the components of vector Z of Brownian motions generating the risks of the economy, be split into the part Z^Q which spans the observations process Q , and the independent part Z^v spanning the unobserved states (that are responsible for the difference between F^Q and F). The dimension of Z^Q must be equal to that of Q . Moreover, the law of motion of Q shall only involve Z^Q , so that it can be written as $dQ = qdt + QdZ^Q$, with a non-singular diffusion matrix Q satisfying regularity conditions needed for the Girsanov theorem to be applicable. Besides, I assume that the last component of Q , i.e. M , is observed through the random component L of the order flow received by the dealer (as defined in Subsection 2.1). This means, among other things, that dL is spanned by the totality of components of dZ^Q in a non-trivial way (i.e. the private order flow reflects the full range of the observed sources of uncertainty in the economy). I will require even more content from this order flow, by making the following assumption about the way a dealer interprets dL .

Assumption 3 The random component dL of the order flow coming from the clients and other dealers has the form

$$dL = g' dQ + g^p dP + g^m dM, \quad (14)$$

where coefficient g^p is positive.

Equation (14) shall be understood as follows. The dealer realizes that the rate of resident dealer purchases of I is linked to the customer purchases from M-resident dealers by the “balance of payments” identity. In view of Assumption c made in Section 3 (zero average customer order flow growth), it is clear that the customer order flow component of dL cannot contain a persistent non-zero trend. The dealer links the balance of payments to Q' and dP/P , and is, therefore, able to establish a one-to-one correspondence between dL and dM . Further, an upward future move in the exchange rate P will induce the others to purchase more (sell less) at his/her own ask (bid) quote rather than losing time by contacting other dealers. This explains the positive sign of g^p .

Now we can formulate the filtering and Bayesian learning process for the dealer.

The vector q of drift coefficients of the observations process Q will be viewed as an Ito process w with the law of motion

$$dw = [a_0 + a_1 w]dt + b dZ^w \quad (15)$$

under the objective probability. All coefficients are bounded \mathbf{F}^Q -adapted processes, which, in addition, may be non-trivial functions of time. This is the state process unobservable directly by the dealer. Instead, each dealer has a belief about w , denoted by \hat{w} , which is a conditional expectation of w given the dealer's current information: $\hat{w}_t = E[w_t | \mathcal{F}_t^Q]$. This is an \mathbf{F}^Q -adapted process, with initial value \hat{w}_0 and the initial variance-covariance matrix \mathbf{w}_0 assumed to be given.

Put $B = \Theta \cdot \Theta^T$. According to the properties of the Kalman-Bucy filter (see Liptser and Shiryaev, 1977, Ch.12, for details), process \hat{w} satisfies the s.d.e.

$$\begin{aligned} d\hat{w}_t &= (a_0 + a_1 \hat{w}_t)dt + \mathbf{w}_t B^{-1} [dQ_t - (a_0 + a_1 \hat{w}_t)dt] \\ &= \{a_0 + a_1 w - (\mathbf{w} \cdot B^{-1} - a_1)(\hat{w} - w)\}dt + \mathbf{w} \cdot (\Theta^{-1})^T dZ^Q \end{aligned} \quad (16)$$

under the objective probability Pr . The variance-covariance process \mathbf{w} satisfies the deterministic Riccati differential equation

$$\frac{d\mathbf{w}}{dt} = a_1 \cdot \mathbf{w} + \mathbf{w} \cdot a_1 + b \cdot b^T - \mathbf{w} \cdot B^{-1} \cdot \mathbf{w} \quad (17)$$

with initial condition \mathbf{w}_0 .

By subtracting (15) from (16), one obtains the s.d.e. for the dealer's estimation error:

$$d(\hat{w} - w) = -(\mathbf{w} \cdot B^{-1} - a_1)(\hat{w} - w)dt + \mathbf{w} \cdot (\Theta^{-1})^T dZ^Q - b dZ^w.$$

This error is a martingale under Pr and, if matrix $\mathbf{w}B^{-1} - a_1$ is stable, then with time, the subjective estimate becomes closer to the true value of w in the mean.

Recall that we assume identical dealers within each nation, and let the size of the \mathbf{M} -resident dealer population be normalized to unity. Then, in equilibrium, the optimal outgoing inter-dealer trade by the domestic dealer, v^i , must satisfy the trivial market clearing condition

$$v_t^i = \hat{m}_t$$

for all t , where the dealer's subjective estimate of the aggregate mean rate of active \mathbf{I} -purchases from non-residents, \hat{m} , is the last component of the conditional expectation \hat{w} of the unobserved state process w (under the given information structure, the dealer solves the optimization problem from Section 2 under probability Pr^* instead of Pr , and all the processes he works with are \mathbf{F}^Q -projections of the objective \mathbf{F} -adapted processes).

In the notations of Section 2, the optimal policy of the dealer implies

$$P = \frac{\mathbf{x}_f X^d}{\mathbf{x}_i X^f} \frac{k^d}{k^f} \frac{1}{k^i} = \mathbf{q} \frac{k^d}{k^f} \frac{1}{k'(v^i)} = \mathbf{q} \frac{k^d}{k^f} \frac{1}{k'(\widehat{m})}. \quad (18)$$

According to the assumptions of Section 3, k^d and k^f are constants. Further, the term \mathbf{q} (which is related to the shadow price parity index Y) is an Ito process whose differential is equal to the total return differential $dR^d - dR^f$ plus a covariance term $c^0 dt$ dependent only on the components of observations covariance matrix \mathbf{Q} . By assuming the linear-quadratic search cost function k given in Subsection 2.1, we are able to write the instantaneous exchange rate change as

$$\widehat{d}P = dR^d - dR^f + c^0 dt - \frac{k_0 d\widehat{m}}{1 + k_0 \widehat{m}} - (\mathbf{u}^d - \mathbf{u}^f) \cdot \frac{k_0 \widehat{b}^m}{1 + k_0 \widehat{m}} dt + \left(\frac{k_0}{1 + k_0 \widehat{m}} \right)^2 |\widehat{b}^m|^2 dt, \quad (19)$$

where \widehat{b}^m is the last row of the diffusion matrix $\mathbf{w} \cdot (\Theta^{-1})^T$ from (16).

Equation (19) describes the generalized uncovered total return parity in an equilibrium with asymmetric dealer information and Bayesian learning from the order flow signal. Here, the first three terms on the right hand side represent the standard UTRP, while the last three terms – a deviation from UTRP, which we call *disparity* (also known as the *country premium* in the exchange rate literature).

Equation (19) allows us to make three general conclusions about the properties of this asymmetric information equilibrium:

1. At times when new information about the aggregate order flow (dM) is being processed by the dealers, the disparity term is a non-trivial seemingly non-stationary process. If the arrival of new information is a one-time event, the drift component of the disparity eventually disappears.
2. The volatilities of both the exchange rate itself and the disparity term in the generalized UTRP equation should be normally higher under asymmetric information, than in a forex market with fully informed dealers.
3. A change in the perceived covariance structure of the drifts of fundamental variables (matrix \mathbf{w}) by the dealers can have a permanent or, at least, very persistent effect on the perceived dynamic of the aggregate order flow drift (process \widehat{m}) and, thereby, on the disparity constant/country premium level. Therefore, the empirically observed revisions of the country premium are explained in the model as a consequence of a revised interpretation of the order flow statistics monitored by the dealers.

Among the asymmetric information equilibria described in this section there are those with sunspot properties, as well. The reason is the self-fulfilling nature of beliefs about process m . The fact can be illustrated by an example where dealers have perfect information about the statistics of all components of \mathbf{Q} except M .

Example: imprecise knowledge of the aggregate order flow only

To treat this case, it is convenient to consider components \mathbf{Q}' as exogenous parameters of the model, so that (redefined) processes $\mathbf{Q} = M$ and $\mathbf{w} = m$ have dimension 1. Covariance parameters $a_0, a_1, b, \mathbf{w}, \mathbf{m} = \mathbf{Q}$ are now scalars. Let the true average trend of the inter-dealer order flow be a martingale given by

$$dm = -a_1 m + b dZ^m,$$

with $a_1 > 0$, i.e. the true m reverts to zero on average. In such an economy, fully informed dealers would have generated a forex market with UTRP satisfied exactly up to an additional random error term dependent on b .

Incompletely informed dealers must learn the true value of m from the order flow observations. The innovation of their beliefs about m is described by the scalar s.d.e.

$$d\hat{m} = -\left[a_1 \hat{m} + \frac{w}{m^2} (\hat{m} - m) \right] dt + \frac{w}{m} dZ^M, \quad (20)$$

and the evolution of the precision of this subjective estimate – by the scalar deterministic differential equation

$$\dot{w} = -\frac{w^2}{m^2} + 2a_1 w + b^2.$$

This covariance converges to the stable steady state $\bar{w} = m^2 \left\{ \left[a_1^2 + \left(\frac{b}{m} \right)^2 \right]^{\frac{1}{2}} - a_1 \right\}$, which

never vanishes unless the diffusion parameter b in the s.d.e. for m is zero. That is, the dynamic of \hat{m} never fully converges to that of m . The coefficient \hat{b}^m in (19) is equal to $\frac{w}{m}$.

Suppose that the true initial value of m is zero, but the subjective initial belief of the representative domestic dealer happens to be $\hat{m}_0 > 0$. Then, from (20) we conclude that $d\hat{m}$ tends to be negative (the exchange rate movements are more downward sloping than what UTRP is prescribing). Equations (18) and (19) in this case tell us that at such times, the dealer's subjective shadow value of the exchange rate is more often higher than the true value, inducing him to initiate I-purchases in the inter-dealer market and, thereby, validate his beliefs. Put differently, the belief that everyone else purchases I, makes him purchase as well (the herding effect). The dynamic of P in (19) is that of the downward adjustment of the exchange rate after the initial “overshooting” move immediately after the formation of the prior belief $\hat{m}_0 > 0$. Besides, on average, $|d\hat{m}| > |dm|$, i.e. the partially informed dealer “overreacts” on the news about the current movement in the FX market, generating a higher volatility registered by outside observers.

The above example showed an extreme case of self-fulfilling beliefs in the inter-dealer market, while, possessing the maximum possible knowledge on other variables, the dealers had no incentives to correct their biased estimates of the aggregate inter-dealer order flow sufficiently. If the estimate update involves other macro state variables as well, one can expect that the sunspot effects in the m -variable will be mitigated.

5. Conclusion

The paper has demonstrated that the long term “macro” factors influencing the exchange rate and the short term information dissemination about these factors among FX dealers and investors can be handled within a common continuous time portfolio optimization model. Properties of equilibria in this model account for

1. fundamental properties such as the uncovered asset return parity of the exchange rate and its generalizations
2. the existence of variable stochastic spreads in a competitive multi-dealer environment
3. concentration of individual dealer ask and bid quotes around a commonly observed clearing price
4. dealer learning about changing fundamentals from private and inter-dealer trades, in the course of which, new information processing can cause deviations from the uncovered parity of total returns for the exchange rate
5. permanent changes in the disparity constant (country premium) as a result of switching between self-fulfilling beliefs about the statistics of the inter-dealer order flow.

As was argued in Section 4, the origin of a change in self-fulfilling beliefs about the inter-dealer order flow can be the arrival of new information about the covariance matrix of the unobserved state process. This information is most likely to improve the currently available one, and only exceptionally, produce totally new pattern of co-movements between the fundamental characteristics of the economy, such as productivity, asset returns or the term structure of interest rates. Thus, if one excludes extreme overhauls of the long-term structural dependencies in the economy, the order flow effects described in the paper are unlikely to have permanent impact on the exchange rate. On the other hand, this impact may be very long-lasting for the reasons of finite speed of learning by the agents, as expressed by the Kalman-Bucy filter equations featured in the text. On the empirical side, the persistence of order flow-caused effects is confirmed by sufficiently long episodes of deviation of the observed exchange rates from the uncovered total return parity.

In spite of the deep insights provided by the inter-temporal dealer optimization model under full rationality, the limits of applicability of this type of models seem to be attained due to high complexity of the solutions. That is, with the exception of a few special cases, the agent is unable to design optimal trading policies in practical terms, even if he is prepared to follow the economist’s recommendations. Therefore, one of the challenges for future research is the adaptation of the presently existing model with full investor-dealer rationality, to a numerically tractable framework of a dynamic system with boundedly rational agents.

Appendix: The Maximum Principle solution of the dealer problem in continuous time

The material of this appendix is based upon Derviz, 1999a.

The general form of the dealer's optimization problem discussed in Section 2 can be symbolically written down as

$$\max_{\ell} E \left[\int_0^T e^{-[J]_t^U} U(X_t; \ell_t) dt + e^{-[J]_T^B} B(T, X_T) \right], \quad (A1)$$

with respect to controls ℓ , subject to the state-transition equation

$$dX = \mathbf{m}(X, \ell)dt + \mathbf{S}(X, \ell)dZ, \quad (\text{A2})$$

the value X_0 of the state process at time $t=0$ given. The state process X is a vector with n components and Z is a vector of d mutually independent standard Brownian motions. B is the so-called final bequest function. Its present value must possess a limit if the time horizon of the optimization problem is infinite ($T = \infty$). Finally, $[\mathbf{J}]_t^s = \int_t^s \mathbf{J}_t^d t$ is the discount factor between periods t and s .

The problem (A1), (A2) can be solved by forming the current value Hamiltonian

$$\mathbf{H}(t, X, \ell, \mathbf{x}, \Xi) = U(X, \ell) + \mathbf{x} \cdot \mathbf{m}(X, \ell) - \text{tr}(\Xi \cdot \mathbf{S}(X, \ell)),$$

which is to be maximized with respect to ℓ_t . Here, \mathbf{x} and \mathbf{X} are the so called *first-* and *second-order adjoint processes* (\mathbf{x} is of the same dimension n as X and \mathbf{X} is an $n \times d$ -matrix), with $\Xi = \mathbf{x} \cdot D_X \mathbf{S}$ along the optimal path. When state X stands for asset holdings, the adjoint process \mathbf{x} can be called the *shadow price* vector of the corresponding group of assets.

Let $[f, g]$ define the predictable co-variation of diffusion processes f and g , and put $d[f, g] = \langle f, g \rangle dt$ (with the standard shorthand $\langle f \rangle$ for $\langle f, f \rangle$). Then the (first-order) adjoint process \mathbf{x} satisfies the stochastic differential equation

$$d\mathbf{x} = \mathbf{x} \cdot (\mathbf{J}_n dt - dA + \langle A \rangle dt) - D_X U dt, \quad (\text{A3})$$

with the $n \times n$ -matrix valued process A defined by

$$dA = D_X \mathbf{m} dt + D_X \mathbf{S} dZ, A_0 = \mathbf{1}_n.$$

The final condition $\mathbf{x}_T = D_X B(T, X_T)$, or an appropriate transversality condition if $T = \infty$, must be added to (A3). The adjoint process \mathbf{x} can be also described as the X -gradient of the value function of the problem (A1), (A2), provided the latter is differentiable.

In the investor-dealer problem of Section 2, the state-transition equation (A2) is linear in the state variable x . Therefore, the coefficient matrix A and its quadratic variation $\langle A \rangle$ for this transition equation are easily seen to be equal to

$$A = \begin{bmatrix} dr^0 + d\Lambda & dr^d & 0 & 0 \\ 0 & -d\mathbf{a}^d & 0 & 0 \\ -\frac{d\Lambda}{P} & 0 & dr^i & dr^f \\ 0 & 0 & 0 & -d\mathbf{a}^f \end{bmatrix},$$

$$\langle A \rangle = \begin{bmatrix} |I^d + \mathbf{I}|^2 & N^d \cdot (I^d + \mathbf{I} - A^d) & 0 & 0 \\ 0 & |A^d|^2 & 0 & 0 \\ -\frac{\mathbf{I}}{P} \cdot (I^d + \mathbf{I} + I^f) & -\frac{\mathbf{I} \cdot N^d}{P} & |I^f|^2 & N^f \cdot (I^f - A^f) \\ 0 & 0 & 0 & |A^f|^2 \end{bmatrix}.$$

The previous two matrix expressions, if substituted into equation (A3), render the adjoint equation system (6) of Section 2.

To derive the expression for the Hamiltonian of the dealer problem, one needs to calculate the terms containing the first and the second order adjoint process. First of all, observe that

$$\mathbf{m} = \begin{bmatrix} x^0(i^d + l) + x^d n^d - \mathbf{d} - X^d k(v^d) - Pk(v^i) + p^a D(p^a) - p^b S(p^b) \\ -x^d a^d + v^d \\ -\frac{x^0 l}{P} + x^i i^f + x^f n^f + v^i - X^f k(v^f) - D(p^a) + S(p^b) \\ -x^f a^f + v^f \end{bmatrix},$$

and the part of the Hamiltonian containing the first order adjoint process is obtained by scalar multiplication of the above column vector by the row vector \mathbf{x} . Further,

$$D_x \mathbf{s} = \begin{bmatrix} I^d + \mathbf{I} & N^d & 0 & 0 \\ 0 & -A^d & 0 & 0 \\ -\frac{\mathbf{I}}{P} & 0 & I^f & N^f \\ 0 & 0 & 0 & -A^f \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} x^0(I^d + \mathbf{I}) + x^d N^d \\ -x^d A^d \\ -\frac{x^0 \mathbf{I}}{P} + x^i I^f + x^f N^f \\ -x^f A^f \end{bmatrix}.$$

Therefore, the second-order adjoint process part of the Hamiltonian is equal to

$$-\left[\mathbf{x}_0, \mathbf{x}_d, \mathbf{x}_i, \mathbf{x}_f\right] \cdot \begin{bmatrix} x^0 |I^d + \mathbf{I}|^2 + x^d N^d \cdot (I^d + \mathbf{I} - A^d) \\ x^d |A^d|^2 \\ -\frac{x^0 (I^d + \mathbf{I} + I^f) \cdot \mathbf{I}}{P} + x^i |I^f|^2 - \frac{x^d N^d \cdot \mathbf{I}}{P} + x^f N^f \cdot (I^f - \mathbf{I}) \\ x^f |A^f|^2 \end{bmatrix},$$

i.e. it does not contain the control variables. In short, maximizing the Hamiltonian with respect to the controls of the dealer problem is equivalent to maximizing the expression

$$\begin{aligned} & \mathbf{x}_0 \left\{ -X^d k(v^d) - Pk(v^i) + p^a D(p^a) - p^b S(p^b) \right\} + \mathbf{x}_i \left\{ v^i - X^f k(v^f) - D(p^a) + S(p^b) \right\} \\ & - \mathbf{x}_0 \mathbf{d} + \mathbf{x}_d v^d + \mathbf{x}_f v^f. \end{aligned}$$

This maximization is fully described by the first order conditions

$$\begin{aligned} u'(\mathbf{d}) &= \mathbf{x}_0, \\ \mathbf{x}_0 P k'(v^i) &= \mathbf{x}_i, \mathbf{x}_0 X^d k'(v^d) = \mathbf{x}_d, \mathbf{x}_i X^f k'(v^f) = \mathbf{x}_f, \\ \mathbf{x}_0 [D + p^a D'] &= \mathbf{x}_i D', \mathbf{x}_0 [S + p^b S'] = \mathbf{x}_i S' \end{aligned}$$

that we use in Subsection 2.2 of the main text.

References

1. Back, K. (1992) Insider trading in continuous time. *Review of Financial Studies* 5, No.3, 387-409.
2. Back, K., H. Cao, and G. Willard (2000) Imperfect competition among informed traders. *Journal of Finance* 55, No.5, 2117-2156.
3. Back, K., and H. Pedersen (1998) Long-lived information and intraday patterns. *Journal of Financial Markets* 1, 385-402.
4. Brennan, M., and H. Cao (1997) International portfolio investment flows. *Journal of Finance* 52, No.5, 1851-1880.
5. Copeland, T., and D. Galai (1983) Information effects and the bid-ask spread. *Journal of Finance* 38, 1457-69.
6. Derviz, A. (1999a) *Adjoint Processes of the Portfolio Optimization Problem and Equilibrium Asset Prices*. Institute of Information Theory and Automation. Research Report No.1954, June.
7. Derviz, A. (1999b) *Generalized Asset Return Parity and the Exchange Rate in a Financially Open Economy*. The Czech National Bank, Monetary Dept., WP. No.12.
8. Detemple, J. (1986) Asset pricing in a production economy with incomplete information. *Journal of Finance* 41, No.2, 383-392.
9. Dixit, A., and R. Pindyck (1994) *Investment under Uncertainty*. Princeton, N.J.: Princeton Univ. Press.
10. Easley, D., and M. O'Hara (1987) Price, trade size, and information in security markets. *Journal of Financial Economics* 19, 69-90.
11. Elliott, R.J. (1982) *Stochastic Calculus and Applications*. Springer-Verlag.
12. Evans, M., and R. Lyons (1999) *Order Flow and Exchange Rate Dynamics*. NBER Working Paper No. 7317, August.
13. Glosten, L. (1989) Insider trading, liquidity, and the role of the monopolist specialist. *Journal of Business* 62, No.2, 211-235.
14. Glosten, L., and P. Milgrom (1985) Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14, 71-100.
15. Goodhart, C., T. Ito, and R. Payne (1996) One day in June 1993: A study of the working of Reuters 2000-2 electronic foreign exchange trading system. In: J. Frankel, G. Galli and A. Giovannini eds. *The Microstructure of Foreign Exchange Markets*. Chicago: Univ. of Chicago Press; 107-182.
16. Ho, T., and H. Stoll (1983) The dynamics of dealer markets under competition. *Journal of Finance* 38, 1053-1075.
17. Kyle, A. (1985) Continuous auctions and insider trading. *Econometrica* 53, 1315-1335.
18. Liptser, R.S., and A.N. Shiryaev (1977) *Statistics of Random Processes, Vol. I and II*. Springer-Verlag.

19. Nadal-De Simone, F., and W.A. Razzak (1999) *Nominal Exchange Rates and Nominal Interest Rate Differentials*, IMF WP/99/141.
20. Perraudin, W., and P. Vitale (1996) Interdealer trade and information flows in a decentralized foreign exchange market. In: J. Frankel, G. Galli and A. Giovannini eds. *The Microstructure of Foreign Exchange Markets*. Chicago: Univ. of Chicago Press.
21. Seasholes, M. (2000) *Smart Foreign Traders in Emerging Markets*, WP, Haas School of Business, University of California, Berkeley (January).
22. Zapatero, F. (1998) Effects of financial innovations on market volatility when beliefs are heterogeneous. *Journal of Economic Dynamics and Control* 22, 597-626.