2003

WORKING PAPER SERIES

/4/

CZECH NATIONAL BANK

WORKING PAPER SERIES

Components of the Czech Koruna Risk Premium in a Multiple-Dealer FX Market

Alexis Derviz

4/2003

THE WORKING PAPER SERIES OF THE CZECH NATIONAL BANK

The Working Paper series of the Czech National Bank (CNB) is intended to disseminate the results of the CNB-coordinated research projects, as well as other research activities of both the staff of the CNB and colaborating outside contributors. This series supersedes previously issued research series (Institute of Economics of the CNB; Monetary Department of the CNB). All Working Papers are refereed internationally and the views expressed therein are those of the authors and do not necessarily reflect the official views of the CNB.

Printed and distributed by the Czech National Bank. The Working Papers are also available at http://www.cnb.cz.

Reviewed by:	Aleš Bulíř	(Czech National Bank)
	Josef Štěpán	(Charles University)
	Geir Høidal Bjønnes	(Stockholm Institute for Financial Research)
Project Coordinator:	Tibor Hlédik	

Components of the Czech Koruna Risk Premium in a Multiple-Dealer FX Market

Alexis Derviz*

Abstract

The paper proposes a continuous time model of an FX market organized as a multiple dealership. The model reflects a number of salient features of the Czech koruna spot market. The dealers have costly access to the best available quotes. They interpret signals from the joint dealer-customer order flow and decide upon their own quotes and trades in the inter-dealer market. Each dealer uses the observed order flow to improve the subjective estimates of the relevant aggregate variables, which are the sources of uncertainty. One of the risk factors is the size of the cross-border dealer transactions in the FX market. These uncertainties have diffusion form and are dealt with according to the principles of portfolio optimization in continuous time. The model is used to explain the country, or risk, premium in the uncovered national return parity equation for the koruna/euro exchange rate. The two country premium terms that I identify in excess of the usual covariance term (a consequence of the "Jensen inequality effect") are: the dealer heterogeneity-induced inter-dealer market order flow component and the dealer Bayesian learning component. As a result, a "dealer-based total return parity" formula links the exchange rate to both the "fundamental" factors represented by the differential of the national asset returns, and the microstructural factors represented by heterogeneous dealer knowledge of the aggregate order flow and the fundamentals. Evidence on the cross-border order flow dependence of the Czech koruna risk premium, in accordance with the model prediction, is documented.

JEL Codes: F31, G11, G29, D49, D82.
Keywords: Bayesian learning, FX microstructure, optimizing dealer, uncovered parity.

---

* The Czech National Bank, Monetary and Statistical Departmant, and Insitute for Information Theory and Automation, Prague, E-mail:Alexis.Derviz@cnb.cz.
Comments and valuable advice by Rich Lyons from the Haas School, U.C. Berkeley, Charles Goodhart from LSE, Carol Osler from the New York FED, and Martin Evans from Georgetown University, are gratefully acknowledged. The usual disclaimer applies.

Nontechnical summary

The ability (also known as uncovered parity) of national asset return differences to reflect exchange rate expectations has proven to be crucially dependent on the assets and market segments used to empirically verify the theoretical result. The disparity between the foreign cash return and the domestic-foreign asset return differential, called the country-, or risk-, premium, turns out to be poorly explained by standard macroeconomic variables. A deeper insight into the asset price formation mechanism is needed to capture the properties of the observed country premium.

In the present paper, the analysis of the risk premium in the CZK/EUR currency pair is undertaken by combining standard international stochastic finance theory with foreign exchange market microstructure methods. The analysis acknowledges the fact that the koruna's exchange rate against the euro does not emerge in an abstract Walrasian market environment. Instead, it is being set by a finite number of market-making dealers who conduct FX operations with customers residing both inside and outside the country. A representative portion of these transactions is executed by domestically licensed dealer banks for non-resident market users. When one takes this microstructural information into account, it becomes possible to uncover the dependence of the seemingly non-stationary risk premium in the CZK/EUR exchange rate on the cross-border order flow in the interbank FX market.

The methodology used leans on the conjecture – supported by informal evidence collected in Czech-resident dealer banks – that an FX dealer objective is a part of the overall portfolio-optimizing behavior of the dealer's bank. Therefore, we model an international portfolio-optimizing market maker whose information, fully in the spirit of microstructure finance theory, comes from the orders placed by his clients. In the course of executing the clients' orders and placing own orders in the inter-dealer market, the dealer learns about both the fundamental properties of the koruna- and euro-denominated assets and the direction in which funds are being moved between them on aggregate. The dealer's information is never perfect, but his rational Bayesian learning leads to a gradual reduction of the error and hence to the convergence of the exchange rate return towards the standard uncovered parity behavior.

The contribution of the paper to the literature consists in demonstrating that:

a) the observed exchange rate properties depend upon the prevailing market microstructure;
b) the flow of FX orders between differently-endowed and informed investors is responsible for the exchange rate *deviation from uncovered parity*, not for the exchange rate as such;
c) FX market makers need information on the aggregate direction of the fund transfer between currencies and do not need to form beliefs about the statistics of the exchange rate itself.

1. Introduction

The paper addresses the issue of modeling the exchange rate risk premium structure in the formula for uncovered parity of national asset returns, for a currency with a known FX-market structure. The model developed for this purpose reflects some salient features of the spot inter-dealer market with the Czech koruna and is applied to the data on its exchange rate against the euro. I demonstrate that the forex microstructure has an impact on the dynamics of the risk premium, by linking the behavior of the latter to

1) the order flow received by domestic resident dealers from other market users,
2) errors in the dealers' assessments of the aggregate cross-border order flow and economic fundamentals.

The constructed continuous time model allows one to address two types of FX-market effects. The first is 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, referring to information extraction procedures and Bayesian belief updating by participants in the inter-dealer market in the face of changing fundamentals, is more short-run in nature. Combining the said objectives, I use the model to derive a "dealer-based" uncovered parity of national asset returns with respect to the expected exchange rate return. This parity theorem contains a country premium term that depends on the order flow from non-resident market users to resident dealers. One term in the premium is present under both perfect and imperfect information and comes from liquidity needs caused by dealer heterogeneity. The other term arises as a consequence of imperfect information of an individual dealer about the aggregate order flow and national asset return statistics. Bayesian learning and belief updating by the dealers then leads to long-lasting shifts in the country premium.

The obtained result generalizes the uncovered parity property of the exchange rate that comes up naturally in any optimizing model of international asset pricing. (Under the Walrasian market clearing assumption, this uncovered parity would follow from the international consumption-based CAPM.) We call it the *Uncovered Total Return Parity* (UTRP), to make a distinction from the much-compromised uncovered interest rate parity of naïve no-arbitrage models. The 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. Accordingly, the return quoted in a secondary market and not the money market loan/deposit rate, which is pre-determined for the time interval in question, constitutes the continuously updated measure of the expected move in the exchange rate (see Derviz, 2002, for details, including an empirical verification of the UTRP). 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 reflect 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 the seemingly unwarranted variability of the country premium in the UTRP formula. The model predicts that deviations from the UTRP should be 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 the UTRP that persist until the new information has been 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 in which the order flow effects are long-lived.

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 usually makes a conceptual distinction between clients and dealers. This happens both in the single-dealer sequential trade models of Glosten and Milgrom, 1985, or Easley and O’Hara, 1987, with a single risk-neutral market maker, and in multiple-dealer models of the Evans and Lyons, 2002, type. This distinction is barely possible in practice. More importantly, the real-life roles of the investor and dealer in the FX market are often combined in the performance of cooperating units within the same company. Most typically, a major commercial or investment bank has dealers who service its customers and participate in the inter-dealer market, but it also operates proprietary trading desks that exercise FX transactions 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 the 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. Accordingly, most of the available price and trade-size data carry no stamp of the purpose of any given FX transaction.

One of the exceptions in the literature in terms of client–dealer separation is the paper by Madhavan and Smidt, 1993, in which a dealer is also an active investor. Similarly to Madhavan and Smidt, I model a synthetic dealer-investor. All agents in the present model are market users. They can be either *local*, trading only with one selected dealer, or *global*, having access to the inter-dealer market. That is, global market users are able to both search for the best dealer quotes and exercise a trade with the dealer of their choice. The only distinction between dealers and the rest of the global market users is that the former assume the market-making function. Apart from that, dealers have the same intrinsic motive to hold foreign currency positions as any other market user, since they are active in the same lines of international business as the latter. The respective roles of local and global market users in the model can be roughly summarized as follows: the order flow of the locals creates noise and can generate exogenous asymmetry in dealer positions; the order flow of the globals acts as a signal that transmits information between dealers.[1]

[1] This understanding of the aggregate signaling effects of the market users’ order flow can be supported by reference to the model of Bernhardt and Hughson, 1997. There, customers split orders between dealers to equalize the marginal benefits from trade at different pricing schedules. The implicit equalization of marginal

The majority of information-oriented forex microstructure models define the signals as unbiased estimates of some "true" or "fundamental" value of the exchange rate. This feature was inherited from microstructure studies of other markets (e.g. equities, where this true value is associated with revealed earnings or dividend data). However, in the case of exchange rates, no such revelation of absolute truth exists. Instead, I let the signals reaching the dealers contain information about the "other side of the market", i.e. the aggregate behavior of the received trades initiators. This allows us to close the model in a natural way. I am modeling neither pure "information" nor pure "inventory" motives for trade, but rather an "information about inventory" motive.

Another highly stylized feature of the existing models of the inter-dealer market is the strict sequencing of information-laden events in them. (Cf. Evans and Lyons, 2002: first a round of dealer trades with customers signaling fundamental information, then an inter-dealer round to redistribute inventory risk between dealers, then another round of customer trades to unwind open positions.) On the contrary, data on the FX trades of a given financial institution reflect three simultaneously evolving random processes pertaining to own investment needs, inter-dealer operations and customer–dealer operations. In other words, at each moment, the dealer observes the three named categories of transactions simultaneously and has to exercise a belief update based on these observations. Under such conditions, a continuous time stochastic model of 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 in continuous time. 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 considers optimizing models for risk-sensitive agents, the stochastic maximum principle becomes an even more powerful tool for analyzing the 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.[2] 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 has a local customer base that trades with him exclusively, there is enough space for a non-trivial bid-ask spread in his quotes. The adverse selection reason for the existence of spreads identified in most models with informationally heterogeneous traders, namely the compensation of losses from trades with the informed by profits made in trades with the

benefits is the reason why, in our model, the global market user order flow observed by one dealer contains information on the order flow received by other dealers as well.

[2] The previously mentioned continuous time models of security trade by Back and colleagues are based on Kyle's auctioneer and do not include dealers explicitly. Moreover, 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 asset value.

uninformed, is implicitly present in this setting as well. The role of the informed is played by the global traders, while that of the uninformed is played 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 generates 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 area of dealer–customer relations.

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. This type of heterogeneity has already found reflection in the 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 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, compared to the majority of domestic investors. (Therefore, cross-border informational asymmetry does not always mean an advantage for residents.) Our model shares the Brennan and Cao view that residency-based informational heterogeneity matters, but also allows for differences in informational endowments between local and global investors, in accordance with the Seasholes conjecture.

The paper is structured as follows. Section 2 summarizes the implications of the general model for the Czech koruna case and provides evidence on the empirical validation of the theoretical result. Section 3 describes the model variables, the structure of uncertainty and the dealer's optimization problem (Subsection 3.1), after which the optimal dealer strategies for passive (at own quotes) and active (at other dealers' quotes) FX trades are derived (Subsection 3.2). (Proofs of the corresponding technical results are given in the Appendix.) I also show how a non-zero bid-ask spread and a competitive mid-quote arise as a result of dealer optimization. Section 4 analyzes the long-run FX transaction price dynamic resulting from the uncovered total return parity. Section 5 discusses order-flow interpretation and updating of beliefs by dealers, leading to deviations from the uncovered parity of returns. Section 6 concludes and indicates possibilities for future research.

2. Summary of the results and their application to the Czech koruna market

2.1 The "dealer-based" uncovered return parity and its implications

Most macroeconomic models of exchange rate formation, if they are microeconomically founded, result in a property of the exchange rate relating it to the difference in the representative returns on domestic and foreign assets:

$$E_t\left[\Delta s_{t+1}\right] = y_{t+1} - y^*_{t+1} + a^0 . \qquad (1)$$

In the above formula, s is the natural logarithm of the nominal exchange rate, y_{t+1} is the domestic asset return between periods t and t+1, and y^*_{t+1} is the foreign asset return between the same dates. E_t denotes the expectation conditioned on public information at date t. This relation is the *Uncovered Total Return Parity* (UTRP, cf. Derviz, 2002) mentioned in the introduction. The UTRP is derived in a Walrasian market with no explicit market structure and no FX trades in equilibrium. If one conducts the analysis to be discussed in Sections 3–5 for a model of a multiple-dealer FX market, the obtained uncovered parity will be different from (1). Although the model to be introduced is in continuous time, it is possible to formulate a discrete time analogue, which can be tested if the appropriate data are available. Specifically, for the model of this paper, one arrives at the equation

$$E_t\left[\Delta s_{t+1}\right] = y_{t+1} - y^*_{t+1} + a^0 + a^1\omega_{t+1}. \qquad (2)$$

Here, ω_{t+1} is the cross-border (i.e. between non-residents and residents) inter-dealer order flow between dates t and t+1. It has a plus sign when non-residents sell korunas against the euro on aggregate, and a minus sign in the opposite case. Equation (2) is the discrete-time counterpart of the *full symmetric information* continuous time *dealer-based uncovered parity* formula (19), to be derived in Section 5. This means that (2) should prevail when all market participants are aware of the correct statistical model for the macroeconomic fundamentals and also know the true law of motion of the aggregate cross-border FX order flow. If information were imperfect or asymmetric, additional terms would appear in (2), corresponding to errors and learning.

The model application requires a specializing assumption on the statistical law of ω. Accordingly, (2) is valid under the assumption that this order flow process is of the Ornstein–Uhlenbeck type (i.e. mean-reverting to zero). This would correspond to non-explosive behavior of the major koruna market participants at times when massive "bandwagon" speculative attacks do not take place (our time span and sample do not include such episodes).

Clearly, the impact of the ω–term in (2) on the exchange rate expectation is intuitively correct: the currency (CZK in this case) is expected to depreciate when non-residents initiate more of its sales against the euro than purchases, and appreciate when they initiate more purchases than sales.

The next subsection gives some empirical evidence on the possibility of explaining the CZK/EUR country premium by cross-border order flow proxies.

2.2 Empirical evidence on the dealer-based uncovered exchange rate parity for the Czech koruna

An interpretation of equation (2) in empirically observable terms would mean that one verifies both the general symmetric information UTRP stated in Section 4 (when (2) is satisfied) and the dealer asymmetry-based explanation of deviations from it discussed in Section 5 (when it is violated). Carrying out the said verification involves:

A. finding the right model for residuals;
B. defining a rule to calculate the average *ex post* exchange rate change over a given period;
C. choosing the representative domestic and foreign assets with returns y and y^*.

Task A presumes discrete time sampling of Itô equation (19), which typically leads to ARMA residuals instead of the desired i.i.d. ones. On the other hand, accommodation of ARMA terms in formal regressions often happens at the cost of reduced explanatory power by the original right-hand-side variables. At present, the best solution seems to be preliminary filtering of the highest frequencies in the exchange rate, the asset return and the aggregate order flow series. These high frequencies correspond to the diffusion terms in the continuous time parity equation.

Task B, in practical terms, means fixing a time horizon over which the uncovered parity will be tested, and a smoothing/averaging procedure over the chosen horizon for the observed exchange rate movement series. As it turns out, the best-performing smoothing horizons vary between one and six months, depending on the analyzed currency pair and the historical period covered by the sample. Once the horizon is picked, the exact choice of averaging procedure does not play a decisive role. At the same time, there are episodes when the UTRP seems to break down for fixed horizons, manifesting itself instead as a co-integration of the exchange rate *level* with the return differential. Differently from outright UTRP-violations, most of which can be explained within the present model, such episodes require the analysis of typical holding periods of the given assets, and are left out of the present discussion.

From the point of view of Task C, long-maturity government bonds proved to be the best instruments for UTRP analysis. The examples given below refer to ten-year bonds as the most widespread category to be found in all the examined economies.

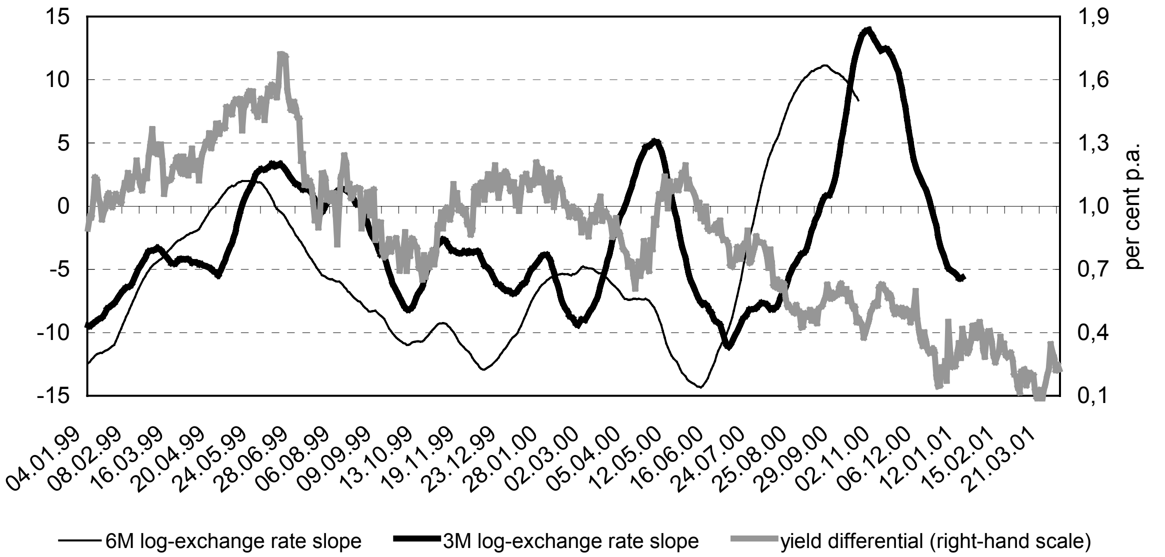
Extensive coverage of both the model and the practical aspects of UTRP testing can be found in Derviz, 2002. Here, to give a general idea of the performance of this concept, I will comment briefly upon the results for two currency pairs: U.S. dollar/euro and Czech koruna/euro. The main finding for both is the very satisfactory performance of the UTRP for the exchange rate changes over the 3M and 6M horizons for a number of sampling periods between 18 and 24 months long on daily data. Figs. 1 and 2 illustrate the results. The *ex post* moving slope of the nominal exchange rate logarithm 3 (6) months ahead of the current date is taken as the smoothing statistic mentioned above.

Even the simple Walrasian symmetric information UTRP stated in (1) can be roughly consistent with the data for as long as one to two years in a row. Longer periods are clearly inappropriate for the UTRP-type reasoning, while the evolution of the disparity (country premium) term in multi-annual samples cannot be ignored. Fig. 3 shows the data for the $/DM exchange rate between Spring 1996 and Spring 2001, 3M exchange rate difference smoothing horizon. The data indicate that there occur regular episodes of disparity term revision. Every such episode is eventually followed by restored validity of the UTRP, but it is impossible to make a single equation such as (2) comply with the whole sample at once.

Our theoretical model predicts that the order flow from non-residents to residents is able to explain a part of the country premium. It is well matched by the data on the cross-border inter-dealer CZK-EUR flows that are available from the balance sheets of the FX dealer banks operating in the Czech koruna market. Fig. 4 shows the disparity term as the difference between the smoothed CZK/EUR rate change (3M smoothing) and the Czech–German government bond differential (i.e. the difference of the two series featured in Fig. 2). This time series is compared to the aggregate net euro-for-koruna purchases of the local Czech dealer banks from the foreign resident partners (clients and dealers). The latter is our proxy of ω. As Fig. 4 demonstrates, the correspondence is at times very close.[3] Another qualitative prediction, backed by the data, refers to the recurring episodes of violation of the standard UTRP. (One such episode concerns the USD/EUR rate starting in Summer 2000, Fig. 1; a number of shorter and less pronounced UTRP-deviations of the CZK/EUR rate are visible in Fig. 2.) The model attributes these violations to dealer-investor learning.

Having provided the rationale behind modeling the CZK country premium with explicit elements of the market microstructure, we now go over to the rigorous definition of the theoretical model and the formal derivation of the quantitative results.

Figure 1: US/EMU 10Y benchmark bond yield differential, 3M- and 6M-moving average of the USD/EUR exchange rate returns



[3] Rigorous econometrics adds little to this outcome of visual inspection and is not commented here.

Figure 2: Czech-German 10Y benchmark bond yield differential, 3M- and 6M-moving average of the CZK/EUR exchange rate returns

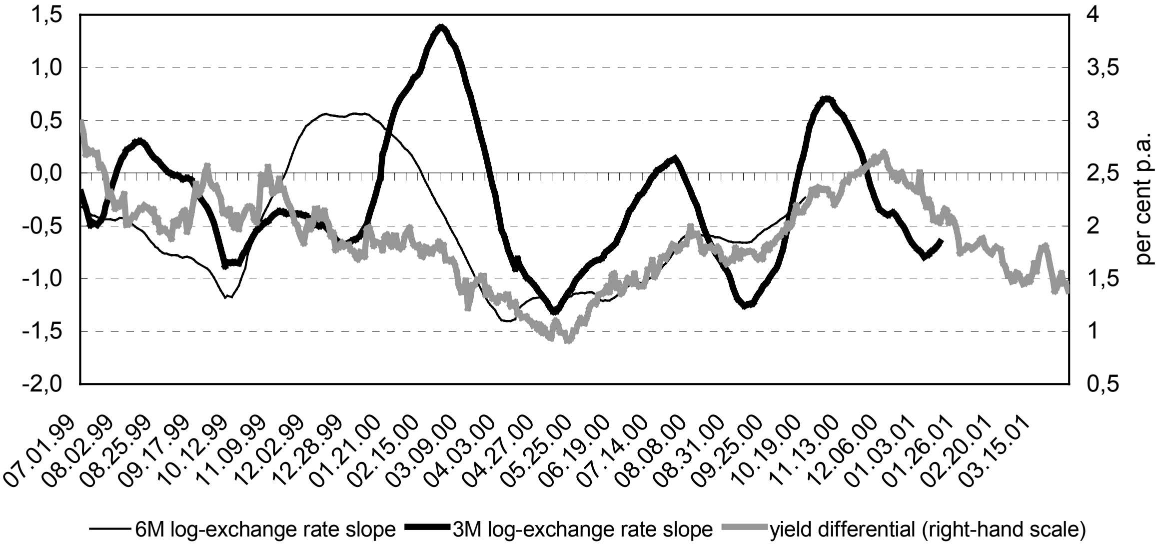


Figure 3: Germany-US 10Y government bond yield differential and the 3M-moving average DEM/USD exchange rate returns, I.1996-III.2001

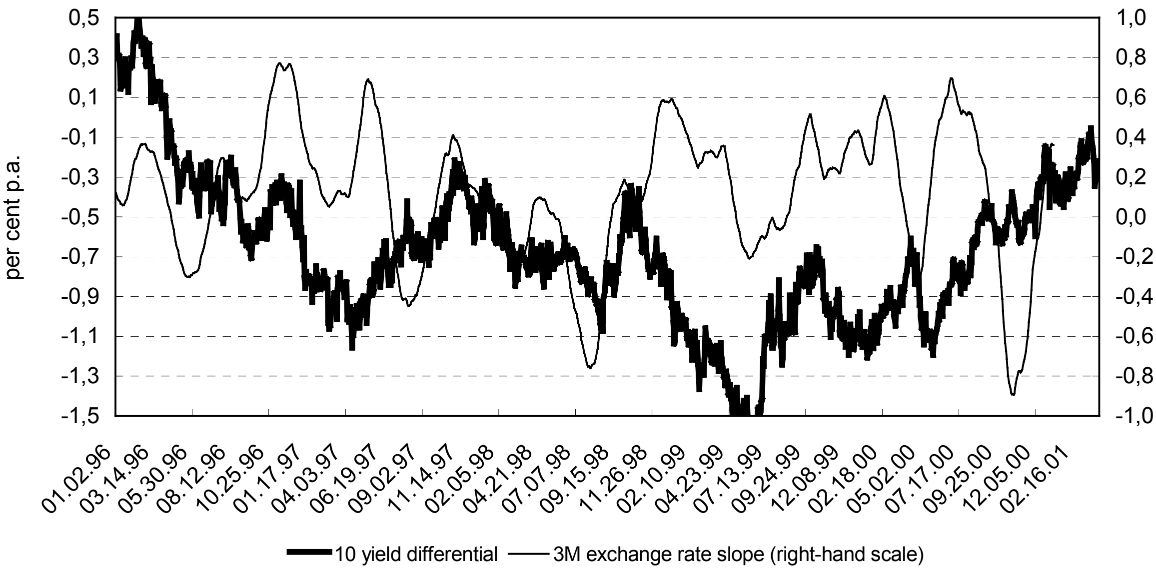
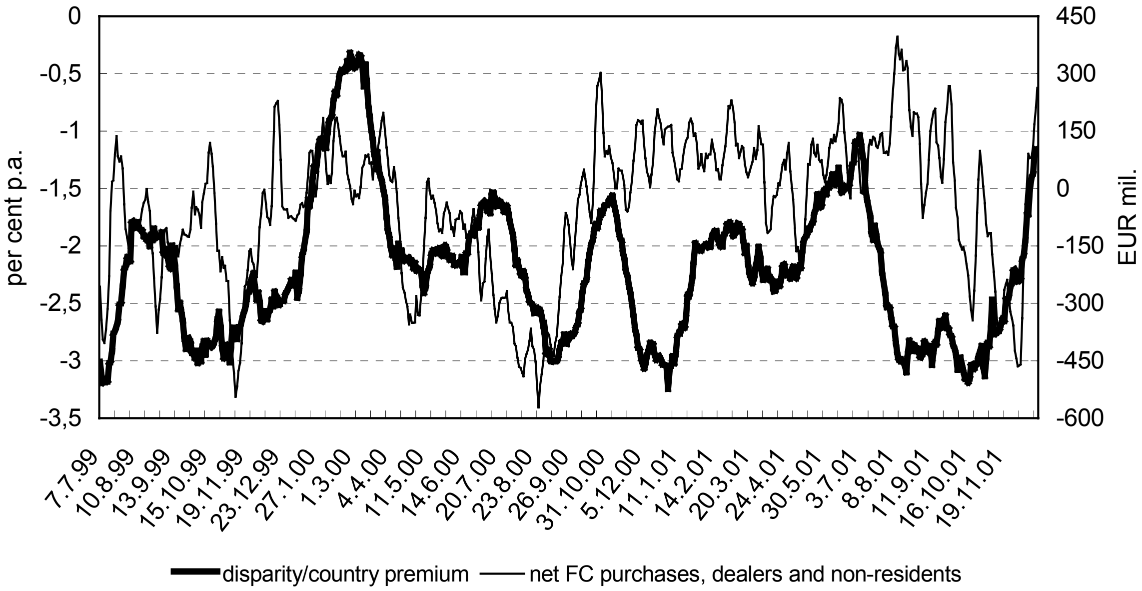


Figure 4: Net cross-border order flow (foreign currency purchases) and deviations from uncovered yield parity, Czech koruna vs. euro, 5D-smoothed OF-series



3. The investor-dealer decision problem

3.1 The model

The world is split into a home country with legal tender M and a foreign country with legal tender I. We take the exchange rate to be the price of I in terms of M. Investors residing in the home country account their operations in M, while those residing in the foreign country use the accounting unit I.[4] 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 FX trades through him at given quotes (they are free to vary the traded quantity from zero to infinity depending on the quotes they see). Alongside them, there are other (global) investors who are able to search for the best quote among all dealers, at a cost. Each dealer also has the ability of costly discovery of the best quote among other dealers. One would expect the investor-dealers to face a lower quote-search cost than investors without a dealing capacity. On the other hand, the non-dealer investors have priority in learning news about fundamentals, to be discussed in Section 5. This balance between the costs and benefits of market making should provide an intuitive justification for the existence of dealers in the model. In the case of a small FX market (e.g. a European currency not belonging to the

[4] The present model accommodates the Jensen inequality-related properties of the expected exchange rate returns (it will be free of the residency-based asymmetry of the latter arising in naïve no-arbitrage settings, also known as Siegel's paradox). See Derviz, 2002, for the details of Siegel's paradox resolution in optimizing models.

eurozone, 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 a source of information update about the statistics of the fundamental variables and the market-wide order flow, read off the received trades. The inter-dealer market is only partially transparent 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 are a part of the overall transaction costs (to be defined below). One can think of an agent with access to electronic and voice brokers who service parts of the market, but who is not sure whether a better price could be found by approaching certain dealers individually. While not modeling these tradeoffs between a broker and individual search explicitly, we adopt the understanding that the best price can always be found, even if the resources dedicated to the search grow faster than the desired transaction amount.

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 $d\alpha^d$. (This is done for the sake of convenience, allowing one to 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, $d\alpha^f$ and X^f respectively (the latter price is in I-units).

Cash holdings x^0 and x^i earn their own rates of return dr^0 and dr^i respectively. In the simplest variant, this 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 Itô process z, dz/z will be shortened to $\widehat{dz}$. The drift and diffusion coefficients of the asset return and price processes introduced above will be denoted as follows:

$$dr^0 = i^d dt + I^d dZ\,,\ dr^i = i^f dt + I^f dZ\,,\ dr^d = n^d dt + N^d dZ\,,\ dr^f = n^f dt + N^f dZ\,,$$
$$d\alpha^d = a^d dt + A^d dZ\,,\ d\alpha^f = a^f dt + A^f dZ\,,$$
$$\widehat{dX}^d = \pi^d dt + \rho^d dZ\,,\ \widehat{dX}^f = \pi^f dt + \rho^f dZ\,.$$

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

The Itô processes r^0, r^d, r^i, r^f, α^d and α^f belong to the exogenous sources of uncertainty in the model. These processes, together with two more to be introduced later in Subsection 2 and Section 5, are assumed to be locally L_2-integrable on $[0,\infty)$ with respect to the physical probability *Pr*, and generate the information filter $\mathbf{F}=(F_t)_{t\geq 0}$. The latter must be augmented by all *Pr*-null sets. All the considered stochastic processes will be adapted to **F** (for every *t*, F_t denotes the partition of the event space corresponding to full information about the operation of the markets; the probability measure *Pr* is defined on F_t for all *t*.) Let the diffusions adapted to **F** be spanned by a vector *Z* of mutually independent standard Brownian motions.

Current FX transaction price

We have assumed that the global market users (dealers or not) are allowed to search in the inter-dealer market for acceptable quotes. Next, we introduce the notion of the best attainable bid and ask prices for a given dealer, P^b and P^a, 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: $P^b= P^a=P$, i.e. the inside spread/touch is zero.

Although, in reality, the touch is positive, it is very small under normal circumstances (i.e. unless the markets become turbulent). Goodhart et al., 1996, observe that interdealer FX spreads on major currency pairs are usually as tiny as 1 pip, or 1/10 of a basis point, for standard amounts. These are the numbers reported by the Reuters D2000-2 electronic brokerage system, i.e., the inside spreads, giving a perfect match with assumption 1 (which just sets the minimal "technical" touch amount to zero for simplicity).

On the theoretical side, 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 the spreads of individual dealers will not fall to zero even when their number becomes large. This is in line with the results to be obtained in the present paper. More generally, the intuitive justification of assumption 1 is related to the information-dissemination ability of the modeled competitive inter-dealer market. 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. There will always be an agent who discovers the deadlocked market and undercuts/overbids 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.[5]

The dealer population is split into sellers and buyers at every given moment. One reason why someone becomes a seller and someone else a buyer is the heterogeneity of asset endowments. 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 domestic asset management, including noisy information on returns, taxation or profit repatriation costs, and vice versa. Domestic dealers may enjoy 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. All these factors can be exacerbated by heterogeneity of beliefs (see Section 5).

The last-observed inter-dealer transaction price P will be identified with the market price. It is the one used in the dealer's accounting. Each agent views it as a strictly positive Itô process.

Investment performance measure

For the later purpose of defining preference over the paths of investor-dealer's actions, we introduce a notional dividend rate δ to be accounted for in every infinitesimal period dt. In the context of high-frequency data (i.e. a very short time step), the defined dividend rate will be understood as an imaginary infinitesimal contribution of the dealer department to the dividend fund of the firm, whose integral, 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), it is usual to define transaction costs that grow in a convex way with respect to transaction volume. It is well known that the 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. In addition, some form of generalized transaction costs seems to lie behind the dealer behavior, even if these costs are not readily measured. In this paper, we take the view that these costs include the cost of a quote search in the inter-dealer market. Specifically, let us assume that in order to buy/sell a volume $v^i dt$ of I at market price P (which is the last observed highest bid if $v^i<0$ and the last observed lowest ask if $v^i>0$), one pays/receives $Pk(v^i)dt$ units of M. Here, k is a strictly increasing and strictly convex function with $k'(0)=1$. A natural example would be a quadratic function of the form

---

[5] Thus, the present model incorporates the observation by Silber, 1984, that dealers are typically competitive on only one side of the market.

$$k(v)=v+k_0v^2/2.$$

For consistency reasons, 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 foreign asset are purchased per period (sold if negative), the M- and I-balances are reduced by amounts $X^d k(v^d)dt$ and $X^f k(v^f)dt$ respectively.

The dealer's order flow

As was mentioned in the introduction, it is difficult to obtain separate data on the customer and inter-dealer components of a dealer's order book. Instead, we distinguish between the local customers of the given dealer (who neither have access to nor exercise FX trades with others) and the participants of the inter-dealer market. Both the global investors who search among available dealer quotes and dealers themselves belong there. Accordingly, the pair of ask and bid I-prices p^a, p^b announced by the dealer will 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, which can go both ways, $x^0 d\Lambda/P=x^0(ldt+\lambda dZ)/P$, generated by Itô process Λ with drift l and diffusion λ. Altogether, the accounted change in the dealer's M-balance, induced by trades at his quotes, is equal to

$$\left(p^a D(p^a)-p^b S(p^b)\right)dt+x^0\left(ldt+\lambda dZ\right)=\left(p^a D(p^a)-p^b S(p^b)\right)dt+x^0 d\Lambda\,, \tag{3}$$

while the corresponding change in the I-balance is

$$\left[S(p^b)-D(p^a)\right]dt-\frac{x^0}{P}\left(ldt+\lambda dZ\right)=\left[S(p^b)-D(p^a)\right]dt-\frac{x^0}{P}d\Lambda\,. \tag{4}$$

Observe, in particular, the last diffusion terms in (3) and (4), coming from the diffusion part of Λ: $x^0\lambda dZ$ and $-\frac{x^0}{P}\lambda dZ$. They reflect the dealer's uncertainty about the order flow from both 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 M-resident, we account for the random order flow in M-terms by expressing it as a stochastic growth rate of the

M-cash account x^0. It can be shown that the I-position in that case evolves exactly as is shown by the last term in equation (2).[6] The presence of the x^0-proportional part $-x^0 \frac{1}{P}$ in the drift component of the order flow is explained by the possibility of probability revision by the dealer (the drift change in Λ in accordance with the Girsanov theorem, see Elliott, 1982). The consequences of such a probability measure change in the case of a Bayesian update of beliefs will be discussed in Section 5.

If the dealer were a foreign resident, the uncertain parts of the order flow that she faces would be xi-proportional, and the certain FX supply- and demand-entries in equations (3), (4), 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=[x0,xd,xi,xf]T of an international investor with the dealer function (with domestic residence):

$$dx^0 = x^0 dr^0 + x^d dr^d - \delta dt - X^d k(v^d)dt - Pk(v^i)dt + \left(p^a D(p^a) - p^b S(p^b)\right)dt$$
$$+ x^0\left(ldt + \lambda dZ\right), \tag{5a}$$
$$dx^d = -x^d d\alpha^d + v^d dt, \tag{5b}$$
$$dx^i = x^i dr^i + x^f dr^f - X^f k(v^f)dt + v^i dt + \left[S(p^b) - D(p^a)\right]dt - \frac{x^0}{P}\left(ldt + \lambda dZ\right), \tag{5c}$$
$$dx^f = -x^f d\alpha^f + v^f dt. \tag{5d}$$

Preferences

The introduced continuous time model is intended to be applicable to high-frequency – such as daily or 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. Earlier, the dividend rate δ in equation (5a) was defined 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 $u(\delta)$ measures the period utility (of the shareholders) derived from receiving contribution δ to the dividend fund. Similarly, the infinitesimal rates dr and $d\alpha$ of asset return and deterioration in the model can be understood as 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 β, as the parameter of a Poisson "death" process. The present model does not require a dealer to close the FX position in a predetermined finite time, nor does it rely on the existence of an exogenous final/underlying value of the currency. Instead, the following construction will replace

[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, for example, Dixit and Pindyck, 1994, Section 3.2.

the artificial shortcut of the "true liquidation value under discovery" often utilized in microstructure models.

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-$e^{\beta dt}$, and liquidate the outstanding engagement in currency I at current prices. The result of liquidation, i.e. the balance x^0+Px^i, 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

$$E\left[\int_t^\infty e^{-\beta(s-t)}\left\{u(\delta_s)+\beta G\left(x_s^0+P_s x_s^i\right)\right\}ds \,\middle|\, F_t\right] \tag{6}$$

with respect to control path $s \mapsto \left(\delta_s, v_s^a, v_s^i, v_s^j, p_s^a, p_s^b\right)$, $s \geq t$, subject to (5), given the current values x_t of asset holdings. Note the appearance of parameter β in front of *G* in the integrand in (4). Since in each period *ds* the firm liquidates with probability 1-$e^{-\beta ds}$, the expected utility derived from the liquidated position is equal to

$$\left(1-e^{-\beta ds}\right)G_s = \frac{1-e^{-\beta ds}}{ds} G_s ds \approx \beta G_s ds\,.$$

If the dealer resides abroad, her liquidation balance entering *G* will be accounted for in I-units and be defined as $\frac{x^0}{P}+x^i$. In either case, we shall assume that *G* decreases to -∞ when cash balances become increasingly negative. Then, the presence of *G* in the period utility prohibits negative cash bubbles of Ponzi type, accomplishing the same objective as the transversality condition of traditional models.

3.2 Optimal policies of the dealer

Following the results of the Appendix, we can characterize the solution to the problem (5), (6) by means of the shadow prices ξ 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:

$$\delta = g(\xi_0)\,, \tag{7a}$$

$$k'(v^i) = \frac{\xi_i}{\xi_0 P}\,, \tag{7b}$$

$$k'(v^d) = \frac{\xi_d}{\xi_0 X^d}\,, \tag{7c}$$

$$k'(v^f) = \frac{\xi_f}{\xi_i X^f}, \tag{7d}$$

$$p^a\left(1-\frac{1}{\varepsilon^a}\right) = \frac{\xi_i}{\xi_0},\ p^b\left(1+\frac{1}{\varepsilon^b}\right) = \frac{\xi_i}{\xi_0}, \tag{7e}$$

where *g* denotes the inverse function to u', $\varepsilon^a = \frac{-p^a D'(p^a)}{D(p^a)}$ is the price elasticity of the local I-demand and $\varepsilon^b = \frac{p^b S'(p^b)}{S(p^b)}$ is the price elasticity of the local I-supply, both observed by the dealer.

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

$$d\xi_0 = \xi_0\left(\beta dt - dr^0 - ldt - \lambda dZ + \left|I^d + \lambda\right|^2 dt\right) + \xi_i\left(\frac{ldt + \lambda dZ}{P} - \frac{\left(I^d + \lambda + I^f\right)\cdot\lambda}{P}dt\right) - \beta G'dt, \tag{8a}$$

$$d\xi_d = \xi_0\left(-dr^d + \left(I^d + \lambda - A^d\right)\cdot N^d dt\right) + \xi_d\left(\beta dt + d\alpha^d + \left|A^d\right|^2 dt\right) - \xi_i \frac{N^d \cdot \lambda}{P} dt, \tag{8b}$$

$$d\xi_i = \xi_i\left(\beta dt - dr^i + \left|I^f\right|^2 dt\right) - \beta PG'dt, \tag{8c}$$

$$d\xi_f = \xi_i\left(-dr^f + \left(I^f - A^f\right)\cdot N^f dt\right) + \xi_f\left(\beta dt + d\alpha^f + \left|A^f\right|^2 dt\right). \tag{8d}$$

If a dealer resides in the foreign country, she faces shadow prices whose laws of motion must be a symmetrically adjusted version of (8), in accordance with her units of account. In other words, in equations (8a)–(8d), one must switch the index pairs (0,*d*) and (*i*,*f*). Also, the term analogous to λdZ and other terms containing λ will appear in equations (7c), (7d) instead of (7a), (7b). Only the quote-setting equation (7e) remains as it is.

Spread

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

$$\frac{p^a}{p^b} = \frac{1+\frac{1}{\varepsilon^a}}{1-\frac{1}{\varepsilon^b}}. \tag{9}$$

If elasticities ε^a and ε^b are high, the log of the right-hand side of (9) is approximately equal to $\frac{1}{\varepsilon^a}+\frac{1}{\varepsilon^b}$. The conjecture about the 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 the inverse elasticities $\frac{1}{\varepsilon^a}$ and $\frac{1}{\varepsilon^b}$ for all customer bases.

Equation (9) would be a standard outcome in any monopolist two-way market maker problem, regardless of the presence of uncertainty. The two less standard elements of the present model are: time-variable and stochastic spreads, and the equilibrium paths of dealer quotes that reflect information about fundamentals, to the extent the latter is disseminated by other parties' actions (see Section 5 below). The latter feature is achieved by considering a competitive inter-dealer market.

The laws of motion (8) of the shadow asset prices characterize them in terms of the *fundamental variables* of the defined economy. These fundamentals in the present model comprise 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 5). The fundamental information accessible to different groups of market participants can have different quality. The possibilities for modeling information dissemination processes with regard to fundamentals within the present approach will be discussed in Section 5. But first, to establish that dealer quotes reflect the processed information on fundamentals competitively (thereby reducing monopolistic welfare loss effects), it is necessary to deduce a number of properties of the possible market equilibria. This is done in the next section.

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

The first-order conditions of optimality (7) can be used to state a fairly general property of the exchange rate dynamics, which we call 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 be obtained 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, 2002, for a model and empirical verification; another approach with more data on leading currencies can be found in Nadal-De Simone and Razzak, 1999.)

To formulate the result, it is necessary to define the total instantaneous rates of return on the domestic and foreign asset. In this model, the return rate is the sum of the instantaneous coupon/dividend rate relative to the current price, dr/X, the capital gain dX/X, less the depreciation rate $d\alpha$:

$$dR^d = \frac{dr^d}{X^d} + \widehat{d}X^d - d\alpha^d = y^d dt + \upsilon^d dZ\,,\; dR^f = \frac{dr^f}{X^f} + \widehat{d}X^f - d\alpha^f = y^f dt + \upsilon^f dZ\,.$$

Next, define the auxiliary variable Y as $Y = \frac{\xi_f X^d}{\xi_d P X^f}$.

Variable Y can be called the *shadow price parity* index. The reason is the fact that, according to Itô's lemma applied to equations (8b) and (8d) for the shadow prices, its stochastic differential satisfies the following property:

$$\widehat{d}Y = dR^d - dR^f - \widehat{d}P + Adt\,. \tag{10}$$

Here, A (the *disparity* term) is the sum of a number of covariance terms (which come about as a consequence of Itô'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 (10) 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 return to the instantaneous return rate differential between representative domestic and foreign securities plus a covariance term:

$$\widehat{d}P = dR^d - dR^f + Adt\,. \tag{11}$$

Note that (11) is symmetric with respect to the residence country of the investor. In other words, this uncovered parity, as opposed to the traditional uncovered interest rate parity, is free of Jensen inequality effects (Siegel's paradox).[7]

A priori, there is no reason to consider 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 (7b)–(7e) the following expression for Y:

$$Y = \frac{k^i k^f}{k^d}\,.$$

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. The possible equilibria would then include ones with constant asset price trends, constant Y and the exchange rate that satisfies the uncovered return parity (11) exactly.

[7] This is true since the shadow price parity index Y is residency-invariant: by switching superscripts d and f and replacing P by $1/P$, one sees that the foreign investor shadow price parity index is equal to $1/Y$. Among other things, the shadow price parity index is constant for foreign investors if and only if it is constant for domestic ones. Recall that parity equation (11) was obtained by Itô-differentiating Y. Therefore, (11) holds regardless of the country of residence, provided one corrects for the additional variance term in the disparity constant A. See Derviz, 2002, for details of the uncovered total return parity model.

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 (5b) and (5d). Then the optimal purchase rates v^d and v^f are positive constants:

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

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

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

Regarding the price processes X^d and X^f, we shall assume that they have no trend ($\pi^d = \pi^f = 0$).

The other three assumptions are aimed at limiting the long-run price trajectories in the forex to the class of bounded ones. That is, attention is restricted to economies and equilibria with the following properties:

a) the exchange rate return drift is bounded from both sides;
b) each of the earlier defined categories of market participants, namely dealers-domestic residents, dealers-foreign residents, global investors and the local customer bases of each dealer, is formed by agents with identical preferences and processing cost functions;
c) the aggregate cumulative client currency supply and demand volumes generated by non-dealer investors in both parts of the world have bounded drifts and are locally riskless (i.e. contain no diffusions).

The benefit of the above restrictions is the existence of common upper and lower bounds for individual inter-dealer orders v^i and, consequently, for individual parameters k^i as well. Assumption b together with the no-diffusion assumption about the non-dealer aggregate I-supplies/demands (the second part of assumption c) will be used in Section 4.

The optimal quote equations (7e), if combined with assumptions a–c, lead to a result about the long-run behavior of the dealer quotes.

Proposition 1 Under assumptions a–c made above, the bid and ask quotes are asymptotically equivalent to

$$\left(1 \pm \frac{1}{\varepsilon^{b,a}}\right)^{-1} YP . \tag{12}$$

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 asymptote (12) itself results from substituting into (7e) the following expression for the shadow cash price ratio ξ_i/ξ_0:

$$\frac{\xi_i}{\xi_0} = Y\frac{k^d}{k^f}P,$$

which follows immediately from (7b)–(7d).

The message of Proposition 1 is the existence of a strong link between the dealer's 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 at which the mid-point between bid and ask is set is pinned down by the dealer's shadow cash prices ratio. The latter is 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-driven (cross-border asset return differential-driven) and is, therefore, competitively determined by all dealers and other investors. Accordingly, no dealer has absolute monopoly power over the clients.

5. 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 the private signals contained in the order flow. In the present model, the Bayesian character of belief updating is implied by the dealer optimization procedure itself, as follows from the Girsanov theorem applied to the probability change in the optimal control of diffusions (see Elliott, 1982, for details). However, an unambiguous prediction of the exchange rate revision direction as a result of a particular news arrival and the corresponding belief update is impossible without the exact specification of the filtering procedure applied by the dealer. In this section, we obtain a more specific characterization of the equilibrium transaction price by making a more specific assumption about the nature of beliefs and signals. As a result, we 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 (5), (6). 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. Symbolically, the relation between Z and Z^* reads

$$dZ=hdt+dZ^*, \tag{13}$$

where h is an **F**-adapted 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 3.1). Every h also generates the trajectories of the dealers' bid and ask quotes and, thereby, the client order flows.

To describe the exact formation mechanism of h, one needs to specify the filtering technique that the agent uses when processing the observations. The most natural choice would be the Kalman–Bucy filter. Equations (15)–(17) below list the basic properties of the dealer's belief updating. The beliefs themselves concern the exogenous variables' drift terms after new information about the current state of the economy has been registered and reflected in the customer order flow.

First, I shall posit the rule of client signal interpretation by the dealer. To do this, I introduce an additional source of exogenous uncertainty into 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

$$dM=mdt+\mu dZ.$$

Given the homogeneity of the national dealer populations (assumption b of Section 4), we conclude that m is formed by the summation of identical active inter-dealer trades v^{i} across the M-resident dealers.

Let us define the *observations process* for the economy as a vector with components – the individual sources of subjectively perceived uncertainty in the model, that is, define the vector diffusion process Q by $Q=[r^{d}, X^{d}, \alpha^{d}, r^{0}, r^{f}, X^{f}, \alpha^{f}, r^{i}, M]^{\mathrm{T}}$. Symbolically, the evolution of Q will be written as $dQ=qdt+\Theta dZ$. This is the dynamic under the objective probability Pr, utilized by those who are able to identify the drift part qdt precisely. (To be able to identify Θ, 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 that 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 the diffusion coefficients Θ. However, dealers have imprecise information about the value of the drift term q. Consequently, an individual dealer's information filter $\mathbf{F}^{Q}$ generated by Q (and completed to satisfy the usual conditions) is cruder than the complete information filter **F** introduced in Section 3. 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 I-currency his/her compatriot population has accumulated on 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 part Z^Q which spans the observations process Q, and the independent part Z^w spanning the unobserved states (responsible for the difference between $\mathbf{F}^Q$ and $\mathbf{F}$). The dimension of Z^Q must be equal to that of Q. Moreover, the law of motion of Q will only involve Z^Q, so that it can be written as $dQ=qdt+\Theta dZ^Q$, with a non-singular diffusion matrix Θ satisfying the regularity conditions needed for the Girsanov theorem to be applicable. Moreover, I assume that the last component of Q, i.e. M, is observed through the random component Λ of the order flow received by the dealer (as defined in Subsection 3.1). This means, among other things, that $d\Lambda$ is spanned by the totality of the components of dZ^Q in a non-trivial way (i.e. the private order flow is disturbed by the full range of the exchange rate-related uncertainty factors). It can be shown that $d\Lambda$ gives the dealer a precise signal about the value of dM (although not about its statistics). Indeed, by summing up transition equations (5c) for all domestic dealers, one arrives at the aggregate relationship

$$\frac{x^{0T}}{P}d\Lambda = dM - \left\{dx^{iT} - \left(x^{iT}dr^i + x^{fT}dr^f - X^f K^f dt + cdt\right)\right\}. \tag{14}$$

Here, x^{0T}, x^{iT} are total M- and I-holdings by the domestic dealer population. K^f is the total net purchase rate of the foreign asset by domestic dealers (the sum of identical k^f terms across the domestic dealer population). Finally, c is the net client sales rate of I to domestic dealers (there are no diffusions in this aggregate rate according to assumption c of Section 4). Invoking assumption b of the same section about investor homogeneity inside categories, we see that the dealer knows both K^f (since he knows his own behavior) and c (which is equal to the sum of the identical terms $S(p^b)$–$D(p^a)$ across the trade partners of all the domestic dealers). This means that $d\Lambda$ indeed signals dM. Equation (14) also shows that the diffusion part of $d\Lambda$ is spanned by the same vector of Brownian motions as that of dQ. That is, the dealer transition equations (5) are consistent with the dealer's information being defined by $\mathbf{F}^Q$, as stated in assumption 2.

Bayesian learning by the dealer

Vector q of drift coefficients of the observation process Q will be viewed as an Itô process w with the law of motion

$$dw = \left[a_0 + a_1 w\right]dt + bdZ^w \tag{15}$$

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 $\widehat{w}$, which is a conditional expectation of w given the dealer's current information: $\widehat{w}_t = E\left[w_t \middle| F_t^Q\right]$. This is an $\mathbf{F}^Q$-adapted process, with initial value $\widehat{w}_0$ and the initial variance-covariance matrix ω_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.

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

under the objective probability *Pr*. The variance-covariance process ω satisfies the deterministic Riccati differential equation

$$\frac{d\omega}{dt} = a_1 \cdot \omega + \omega \cdot a_1 + b \cdot b^T - \omega \cdot B^{-1} \cdot \omega \tag{17}$$

with initial condition ω_0.

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

$$d(\hat{w} - w) = -(\omega \cdot B^{-1} - a_1)(\hat{w} - w)dt + \omega \cdot (\Theta^{-1})^T dZ^Q - b dZ^w . \tag{18}$$

This error is a martingale under *Pr* and, if matrix $\omega B^{-1} - a_1$ 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 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 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 4 under probability Pr^* instead of *Pr*, and all the processes he works with are $\mathbf{F}^Q$-projections of the objective **F**-adapted processes).

In the notations of Section 4, the optimal policy of the dealer implies

$$P = \frac{\xi_f X^d}{\xi_i X^f} \frac{k^d}{k^f} \frac{1}{k^i} = \theta \frac{k^d}{k^f} \frac{1}{k'(v^i)} = \theta \frac{k^d}{k^f} \frac{1}{k'(\hat{m})} .$$

According to the assumptions of Section 4, k^d and k^f are constants. Further, the term θ (which is related to the shadow price parity index Y) is an Itô process with differential 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 Θ. By assuming the linear-quadratic search cost function k given in Subsection 3.1, we are able to formulate the principle result of the paper.

Proposition 2 (*Dealer-based uncovered total return parity*) *The instantaneous exchange rate return in an inter-dealer FX market with imperfect dealer information about the drift terms of the fundamentals is equal to*

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

where $\widehat{b}^m$ is the last row of the diffusion matrix $\omega\cdot\left(\Theta^{-1}\right)^T$ 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 are a deviation from the UTRP, which we call *disparity* (also known as the *country premium* in international finance). The main difference compared to standard uncovered parity formulae is the presence of the term containing the dealer-perceived trend of the cross-border order flow, $d\widehat{m}$. The latter can be split into dm, the full-information law of motion of m, and the error term $d(\widehat{m}-m)$, given by the last row of (18).

Even under full symmetric information about the cross-border order flow statistics, (19) shows why the exchange rate behavior may frequently deviate from the uncovered parity rule. Any liquidity-induced movement in the aggregate order flow trend m leads to an additional component in the observed country premium. For instance, if non-residents accelerate their purchases of I, generating a positive shift of m, the I-currency appreciates more than prescribed by the uncovered parity. When, in addition, the dealers are imperfectly informed about m, even the aforementioned liquidity-adjusted UTRP is violated (see the example below). Formally, the term in (19) containing $d(\widehat{m}-m)$ creates an additional individually perfectly rational disparity. Both effects would be impossible in a purely Walrasian FX market.

The dealer-based generalized UTRP (19) indicates that the non-stationary disparity term behavior corresponds to a period when the aggregate cross-border inter-dealer order flow dM has a non-zero time-dependent drift parameter m. By collecting all the diffusion parameter terms in a single (close to constant) term c^1, one can rewrite (19) as

$$\widehat{dP} = dR^d - dR^f + c^1 dt - \frac{k_0 d\widehat{m}}{1+k_0\widehat{m}}. \quad (20)$$

Even if one ignores the difference between objective m and the dealer's subjective $\widehat{m}$, a non-zero aggregate I-purchase or I-sale pressure resulting in a non-trivial dynamic of m is able to generate a temporary UTRP-deviation. Let us assume that the true law of motion of m is given by the Ornstein–Uhlenbeck equation

$$dm = -a^m m dt + bdZ^m , \tag{21}$$

with $a^m>0$, i.e. m reverts to zero on average. Then (20) renders a perfectly intuitive result of the domestic currency (M) depreciating *relative to the national asset return differential* when residents give up M in favor of I (m positive), and appreciating relative to this differential when residents give up I and accumulate M (m negative).

The cross-border inter-dealer order flow is able to explain UTRP-deviations in both the symmetric and asymmetric information cases. An important distinction of the asymmetric information case is that, differently from (21), the subjective drift parameter $\hat{m}$ is no longer a martingale. This is why it is more natural to associate the persistence of the non-stationary disparity term with the dealers' Bayesian information-acquisition at a finite speed, as described by (16), (17).

The remaining disparity not explained by the combination of (20) and (21) can be conjectured to come from the asymmetric information phenomena outlined at the beginning of this section. In this case, (19) must be invoked in full generality, including the difference between m and $\hat{m}$. Then the corresponding dynamic equilibrium becomes explicitly dependent on the dealers' learning process. One comes to three general conclusions about the properties of these asymmetric information equilibria:

a) At times when new information about the aggregate order flow (dM) is being processed by the dealers, the disparity term is a non-trivial and seemingly non-stationary process. If the arrival of new information is a one-time event, the drift component of the disparity eventually disappears.
b) The volatilities of both the exchange rate itself and the disparity term in the generalized UTRP equation should be higher under asymmetric information than in a forex market with fully informed dealers.
c) The arrival of a new order flow signal induces a change in a dealer's perceived covariance structure of the drifts of fundamental variables (matrix ω). Therefore, new information can have a permanent or, at least, very long-lived effect on the perceived dynamic of the aggregate order flow drift (process $\hat{m}$) and, thereby, on the disparity/country premium level. 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 many with sunspot properties. The reason is the self-fulfilling nature of the beliefs about process m. This fact can be illustrated by an example where dealers have perfect information about the statistics of all components of Q except M.

Example: imprecise knowledge of the aggregate order flow only

To treat this case, it is convenient to consider components Q' as exogenous parameters of the model, so that (redefined) processes $Q=M$ and $w=m$ have dimension 1. Covariance parameters a_0, a_1, b, ω, $\mu=\Theta$ are now scalars. Let the true average trend of the inter-dealer order flow be a martingale of the (21) type. In such an economy, fully informed dealers would have generated a forex market with the 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.

$$d\hat{m} = -\left[a_1 \hat{m} + \frac{\omega}{\mu^2}(\hat{m} - m) \right] dt + \frac{\omega}{\mu} dZ^M , \tag{22}$$

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

$$\dot{\omega} = -\frac{\omega^2}{\mu^2} + 2a_1\omega + b^2 .$$

This covariance converges to the stable steady state $\overline{\omega} = \mu^2 \left\{ \left[a_1^{\,2} + \left(\frac{b}{\mu} \right)^2 \right]^{\frac{1}{2}} - a_1 \right\}$, 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 $\frac{\omega}{\mu}$.

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 (22) we conclude that $d\hat{m}$ tends to be negative (the exchange rate movements are more downward sloping than prescribed by the UTRP). 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 is purchasing I makes him purchase it 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 $\hat{m}_0 > 0$. In addition, on average, $|d\hat{m}| > |dm|$, i.e. the partially informed dealer "overreacts" to the news about the current movement in the FX market, generating a higher volatility registered by outside observers.

The above example dealt with an extreme case of self-fulfilling beliefs in the inter-dealer market, while the dealers, possessing the maximum possible knowledge of other variables, had no incentive 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.

6. Conclusion

This paper has developed a model of optimizing FX dealers who use the received order flow to improve their knowledge of fundamentals in a Bayesian manner. The model reflects a number of typical features of spot trading in the Czech koruna. It combines the traditional logic of international macroeconomics with techniques of continuous time finance and forex microstructure theory. The latter originally declared the intention to deal with exchange rate formation in terms that could be recognized by financial market practitioners, as opposed to textbook macroeconomic lessons, which are rarely reflected accurately by FX traders. The problem is especially urgent in the eyes of the 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. However, the best-known microstructure models currently available operate with conceptual shortcuts and information-theoretic constructions that make their messages even more, not less, distant from the dealer-room language than the standard propositions of classroom macroeconomics.[8] An additional problem is to derive from any of these models an empirically meaningful corollary testable on the available data. The present paper makes a step in this direction by offering evidence that the cross-border order flow in the Czech koruna/euro spot market is able to explain a part of the CZK risk premium. This might be a useful lesson for any macro model of the Czech economy containing an exchange rate equation with a risk premium. Indeed, certain entries in the balance of payments financial account, serving as a proxy for the FX order flow from non-residents to residents, can help to explain the observed deviations of the exchange rate from the uncovered parity trajectory.

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

1) 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) the 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.

[8] This is the author's experience from dealing room conversations in a number of commercial and reserve banks. The conclusion is, not surprisingly, that dealers are usually well versed in the uncovered interest rate parity argument (and even know that it does not hold). They can also easily comment on the signal impact of fresh data about inflation or GDP on the exchange rate. It is, however, totally unrealistic to expect a dealer to analyze the separating-versus-pooling nature of equilibrium in the 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.

As was argued in Section 5, a change in self-fulfilling beliefs about the inter-dealer order flow can result from the arrival of new information about the drift and the covariance matrix of the unobserved state process. This information is most likely to improve the currently available one, and only exceptionally produce a totally new pattern of co-movements between the fundamental characteristics of the economy such as productivity, asset returns and 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 this paper are unlikely to have a permanent impact on the exchange rate. On the other hand, this impact may be very long-lasting, owing to the finite speed of learning by the agents. 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.

References

BACK, K. (1992): "Insider Trading in Continuous Time", *Review of Financial Studies* 5, No.3, 387–409.

BACK, K., CAO, H. AND WILLARD, G. (2000): "Imperfect Competition Among Informed Traders", *Journal of Finance* 55, No. 5, 2117–2156.

BACK, K., AND PEDERSEN, H. (1998): "Long-lived Information and Intraday Patterns", *Journal of Financial Markets* 1, 385–402.

BERNHARDT, D., AND HUGHSON, E. (1997): "Splitting Orders", *Review of Financial Studies* 10, No. 1, 69–101.

BISMUT, J.M. (1976): "Théorie Probabiliste du Contrôle des Diffusions", *Memoirs of the American Mathematical Society* 4, No. 167.

BRENNAN, M., AND CAO, H. (1997): "International Portfolio Investment Flows", *Journal of Finance* 52, No. 5, 1851–1880.

CADENILLAS, A., AND KARATZAS, I. (1995): "The Stochastic Maximum Principle for Linear, Convex Optimal Control with Random Coefficients", *SIAM Journal of Control and Optimization* 33, No. 2, 590–624.

COPELAND, T., AND GALAI, D. (1983): "Information Effects and the Bid-Ask Spread", *Journal of Finance* 38, 1457–1469.

DERVIZ, A. (2002): "The Uncovered Parity Properties of the Czech Koruna", *Prague Economic Papers* 16, No. 1, 17–39.

DETEMPLE, J. (1986): "Asset Pricing in a Production Economy with Incomplete Information", *Journal of Finance* 41, No. 2, 383–392.

DIXIT, A., AND PINDYCK, R. (1994): *Investment under Uncertainty,* Princeton, N. J.: Princeton Univ. Press.

EASLEY, D., AND O'HARA, M. (1987): "Price, Trade Size, and Information in Security Markets", *Journal of Financial Economics* 19, 69–90.

ELLIOTT, R.J. (1982): *Stochastic Calculus and Applications,* Springer–Verlag.

EVANS, M., AND LYONS, R. (2002): "Order Flow and Exchange Rate Dynamics", *Journal of Political Economy* 110 (February), 170–180.

GLOSTEN, L. (1989): "Insider Trading, Liquidity, and the Role of the Monopolist Specialist", *Journal of Business* 62, No. 2, 211–235.

GLOSTEN, L., AND MILGROM, P. (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics* 14, 71–100.

GOODHART, C., ITO, T., AND PAYNE, R. (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.

HAUSMANN, U.G. (1981): "On the Adjoint Process for Optimal Control of Diffusion Processes", *SIAM Journal of Control and Optimization* 25, 1557–1586.

HO, T., AND STOLL, H. (1983): “The Dynamics of Dealer Markets under Competition”, *Journal of Finance* 38, 1053–1075.

KYLE, A. (1985): “Continuous Auctions and Insider Trading”, *Econometrica* 53, 1315–1335.

LIPTSER, R.S., AND SHIRYAEV, A.N. (1977): *Statistics of Random Processes, Vol. I and II*. Springer–Verlag.

MADHAVAN, A., AND SMIDT, S. (1993): An Analysis of Changes in Specialist Inventories and Quotations”, *Journal of Finance* 48, 1595–1628.

NADAL-DE SIMONE, F., AND RAZZAK, W.A. (1999): *Nominal Exchange Rates and Nominal Interest Rate Differentials*, IMF WP/99/141.

PENG, S. (1990): “A General Stochastic Maximum Principle for Optimal Control Problems”, *SIAM Journal of Control and Optimization* 28, No. 4, 966–979.

PERRAUDIN, W., AND VITALE, P. (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.

SEASHOLES, M. (2000): *Smart Foreign Traders in Emerging Markets*, WP, Haas School of Business, University of California, Berkeley (January).

SILBER, W.L. (1984): “Marketmaker Behavior in an Auction Market: An Analysis of Scalpers in Future Markets”, *Journal of Finance* 39, 937–953.

ZAPATERO, F. (1998): “Effects of Financial Innovations on Market Volatility when Beliefs are Heterogeneous”, *Journal of Economic Dynamics and Control* 22, 597–626.

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

The following results utilize the adjoint equation techniques in stochastic optimal control of diffusions pioneered by Bismut, 1976, and Hausmann, 1981, and further developed in Peng, 1990, and Cadenillas and Karatzas, 1995.

The dealer's optimization problem discussed in Section 3 can be symbolically written down as

$$\max_{\ell} E\left[\int_0^T e^{-[\vartheta]_0^t} U(X_t;\ell_t)dt + e^{-[\vartheta]_0^T} B(T,X_T)\right], \tag{A1}$$

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

$$dX = \mu(X,\ell)dt + \sigma(X,\ell)dZ\,, \tag{A2}$$

the value X_0 of the state process at time t=0 being 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, $[\vartheta]_t^s = \int_t^s \vartheta_\tau d\tau$ is the discount factor between periods t and s.

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

$$\mathsf{H}(t,X,\ell,\xi,\Xi) = U(X,\ell) + \xi\cdot\mu(X,\ell) - tr(\Xi\cdot\sigma(X,\ell)),$$

which is to be maximized with respect to ℓ_t. Here, ξ and Ξ are the *first-* and *second-order adjoint processes* (ξ is of the same dimension n as X and Ξ is an nxd-matrix), with $\Xi = \xi\cdot D_x\sigma$ along the optimal path. When state X stands for asset holdings, the adjoint process ξ 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 $d\lfloor f,g\rfloor = \langle f,g\rangle dt$ (with the standard shorthand $\langle f\rangle$ for $\langle f,f\rangle$). Then the (first-order) adjoint process ξ satisfies the stochastic differential equation

$$d\xi = \xi\cdot\left(\vartheta\mathbf{1_n}dt - dA + \langle A\rangle dt\right) - D_X U dt\,, \tag{A3}$$

with the *nxn*-matrix valued process *A* defined by

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

The final condition $\xi_T = D_X B(T, X_T)$, or an appropriate transversality condition if $T = \infty$, must be added to (A3). The adjoint process ξ can also be 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

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

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

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 second-order adjoint process. First of all, observe that

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

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 ξ. Further,

$$D_x\sigma = \begin{bmatrix} I^d+\lambda & N^d & 0 & 0 \\ 0 & -A^d & 0 & 0 \\ -\dfrac{\lambda}{P} & 0 & I^f & N^f \\ 0 & 0 & 0 & -A^f \end{bmatrix}, \quad \sigma = \begin{bmatrix} x^0(I^d+\lambda)+x^dN^d \\ -x^dA^d \\ -\dfrac{x^0\lambda}{P}+x^iI^f+x^fN^f \\ -x^fA^f \end{bmatrix}.$$

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

$$-[\xi_0,\xi_d,\xi_i,\xi_f]\cdot\begin{bmatrix} x^0\left|I^d+\lambda\right|^2+x^dN^d\cdot\left(I^d+\lambda-A^d\right) \\ x^d\left|A^d\right|^2 \\ -\dfrac{x^0(I^d+\lambda+I^f)\cdot\lambda}{P}+x^i\left|I^f\right|^2-\dfrac{x^dN^d\cdot\lambda}{P}+x^fN^f\cdot\left(I^f-\lambda\right) \\ x^f\left|A^f\right|^2 \end{bmatrix},$$

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

$$\xi_0\left\{-X^dk(v^d)-Pk(v^i)+p^aD(p^a)-p^bS(p^b)\right\}+\xi_i\left\{v^i-X^fk(v^f)-D(p^a)+S(p^b)\right\}$$
$$-\xi_0\delta+\xi_dv^d+\xi_fv^f.$$

This maximization is fully described by the first-order conditions

$$u'(\delta)=\xi_0,$$
$$\xi_0Pk'(v^i)=\xi_i,\ \xi_0X^dk'(v^d)=\xi_d,\ \xi_iX^fk'(v^f)=\xi_f,$$
$$\xi_0\left[D+p^aD'\right]=\xi_iD',\ \xi_0\left[S+p^bS'\right]=\xi_iS',$$

that we use in Subsection 4.2 of the main text.

CNB Working Papers Series

4/2003	Alexis Derviz:	*Components of the Czech Koruna Risk Premium in a Multiple-Dealer FX Market*
3/2003	Vladimír Benáček: Ladislav Prokop Jan Á. Víšek	*Determining Factors of the Czech Foreign Trade Balance: Structural Issues in Trade Creation*
2/2003	Martin Čihák: Tomáš Holub	*Price convergence to the EU: What do the 1999 ICP data tell us?*
1/2003	Kamil Galuščák: Daniel Münich	*Microfoundations of the wage inflation in the Czech Republic*
4/2002	Vladislav Flek: Lenka Marková Jiří Podpiera	*Sectoral productivity and real exchange rate appreciation: Much ado about nothing?*
3/2002	Kateřina Šmídková: Ray Barrell Dawn Holland	*Estimates of fundamental real exchange rates for the five EU pre-accession countries*
2/2002	Martin Hlušek:	*Estimating market probabilities of future interest rate changes*
1/2002	Viktor Kotlán:	*Monetary policy and the term spread in a macro model of a small open economy*

Czech National Bank
Economic Research Department
Na Příkopě 28, 115 03 Praha 1
Czech Republic
phone: +420 2 244 12 321
fax: +420 2 244 14 278
e-mail: eva.grenarova@cnb.cz