

# Inventory Information

H. Henry Cao\*

Martin D.D. Evans

Richard K. Lyons

21 October 2003

## Abstract

When information about fundamentals is symmetric, can information-based trade still arise? Consider bond and FX markets, where private information about nominal cash flows is generally absent, but participants are convinced that superior information exists. We analyze a class of asymmetric information—inventory information—that is unrelated to fundamentals, but still forecasts future price (by forecasting future discount factors). We find that inventory information in FX does indeed forecast discount factors, and does so over both short and long horizons. The permanent effect from inventory information ranges between 15 and 30 percent of that from public information.

## Correspondence

Professor Richard K. Lyons  
Haas School of Business, UC Berkeley  
Berkeley, CA 94720-1900  
Tel: 510-642-1059, Fax: 510-643-1420  
Email: [lyons@haas.berkeley.edu](mailto:lyons@haas.berkeley.edu)  
Web: [faculty.haas.berkeley.edu/lyons](http://faculty.haas.berkeley.edu/lyons)

---

\* Forthcoming, *J. of Business*. Respective affiliations are Kenan-Flagler School of Business ([caoh@bschool.unc.edu](mailto:caoh@bschool.unc.edu)), Georgetown University and NBER ([evansm1@georgetown.edu](mailto:evansm1@georgetown.edu)), and UC Berkeley and NBER ([lyons@haas.berkeley.edu](mailto:lyons@haas.berkeley.edu)). We thank the following for valuable comments: an anonymous referee, Bruno Biais, Doug Breedon, Jennifer Conrad, Jennifer Carpenter, Chris D'Souza, Rick Green, Frank Hatheway, David Hirshleifer, Burton Hollifield, Jonathan Karpoff, Ananth Madhavan, Michael Melvin, Maureen O'Hara, Steve Slezak, James Wang, Ingrid Werner, and seminar participants at CMU, Duke, Harvard, Washington, Ohio State, USC, UNC, Penn State, Florida, Arizona State, UC Riverside, Insead, Oslo, Copenhagen, HEC, Toulouse, and the 1999 WFA. Evans and Lyons thank the National Science Foundation for financial assistance.

## Inventory Information

When information about fundamentals is symmetric, can information-based trade still arise? The question is central to our understanding of some important markets, e.g., government bond and FX markets, where the premise of symmetric information about fundamentals is natural, yet participants are convinced that superior information exists (see, e.g., Cheung and Wong 2000, Ciccotello and Hatheway, 2000, Massa and Simonov 2003, Green 2002). To resolve the puzzle, we address a class of information that we call inventory information. Inventory information is orthogonal to fundamentals (that is, orthogonal to cash flows from holding the asset, e.g., future coupons and principal in the case of a bond, and future interest differentials in the case of FX). At the same time, inventory information has the power to forecast future price. It does so by forecasting future discount factors (over both short and long horizons).<sup>1</sup> Inventory information is also asymmetric: initially it manifests in marketmaker positions, which are privately observed, and is only subsequently impounded in price. These two features—ability to forecast price and private observation—provide the basis for information-based trade.

Beyond clarifying this information-based trade, our analysis shows that price effects from non-fundamental trades are of three distinct types. The first is idiosyncratic to individual marketmaker prices and is transitory—the so-called inventory effects from microstructure theory. The second type is common to all marketmaker prices (intraday risk premia) and is also transitory. In traditional models with a single marketmaker, the first and second of these are effectively the same. The third type of price effect from non-fundamental trade is common to all marketmaker prices and is permanent. This third price effect persists even after marketmaker inventories are fully shared economy-wide (e.g., permanent portfolio effects across assets that are imperfect substitutes; see, e.g., Scholes 1972 and Branson and Henderson 1985). Traditional models of securities trading do not address this third effect: in most models, inventory risk is assumed diversifiable at the economy-wide level.<sup>2</sup> Our empirical methodology is designed to accommodate this third, permanent effect.

Our empirical strategy for isolating price effects from inventory information is based on the FX market, a market where superior information about future cash flows is generally absent.

---

<sup>1</sup> The asset pricing literature addresses permanent price effects from changing discount factors, but without the information-based trading that is our focus here. For example, Duffie and Constantinides (1996) show that permanent shocks to endowments cannot be insured against and therefore affect prices. That permanent shocks to endowments are relevant empirically is shown by Storesletten, Telmer and Yaron (2003).

<sup>2</sup> This third type of effect also arises in the models of Lyons (1997) and Saar (1999) and is addressed empirically in equity markets by Kaul, Mehrotra, and Morck (2000) and Wurgler and Zhuravskaya (2002).

The equity market does not share this feature of FX. For equities, empirical findings that order flow induces permanent price changes are generally interpreted as evidence that initiating counterparties have superior information about future cash flows. This interpretation is sensible because (i) many factors affecting equity cash flows—e.g., earnings—are well understood, and (ii) it is credible that some traders have superior information about these factors. Neither condition applies to the FX market. The factors affecting cash flows to holding major currencies are, empirically, unknown: movements in spot rates are almost completely disconnected from the macroeconomic fundamentals that determine these cash flows (Meese and Rogoff 1983, Frankel and Rose 1995).<sup>3</sup> Moreover, even in the few cases where macro news events do have significant (but small) effects on spot rates (Andersen et al. 2003), there is no evidence that individuals have access to information prior to scheduled public release. In fact, the only instance where asymmetric information about future cash flows might become an issue in FX is when a central bank intervenes. However, our sample period contains no interventions, so even this potential source of trade-related cash-flow information is not present. (In any event, most interventions in FX are “sterilized” as a matter of course, meaning that their effects on macroeconomic variables like interest rates and money supplies are neutralized.) In sum, the empirical case that FX trades do not contain information about cash flows is a strong one.

We find that price effects from inventory information are large and include both transitory and permanent components. The impact effect of inventory information is roughly 50 percent of that from public news (the latter being the whole story under symmetric information). After 30 minutes, the transitory effect of inventory information dies out, and prices reflect a permanent effect between 15 and 30 percent of that from public news (depending on model variations).<sup>4</sup> These permanent effects are precisely estimated and economically significant. They imply that \$1 billion increase in unanticipated dollar purchases increases the deutsche mark price of a dollar by 0.17 to 0.30 percent forever.

Implications of our results extend beyond microstructure finance, touching on exchange rate economics and asset pricing more broadly. Addressing microstructure first, our results suggest that information effects detected in past empirical work may have been interpreted mistakenly as information about fundamentals, whereas in reality some information effects reflect information about discount rates. (Consider, for example, past empirical work based on the

---

<sup>3</sup> As Meese (1990) put it in his survey, “the proportion of (monthly or quarterly) exchange rate changes that current models can explain is essentially zero”. This is echoed in the later survey by Frankel and Rose (1995): “To repeat a central fact of life, there is remarkably little evidence that macroeconomic variables have consistent strong effects on floating exchange rates”. The consensus remains true today.

<sup>4</sup> Transitory inventory effects on price are documented in Lyons (1995), Hansch et al. (1998), Madhavan and Sofianos (1999), and Reiss and Werner (1999).

methodology of Hasbrouck 1991.)<sup>5</sup> Turning to exchange rate economics, the results of this paper speak directly to two big debates. The first is the degree to which order-flow effects on price persist. That flow effects on price exist, and account for a large share of short horizon return variation (e.g., daily) is well established (see the survey in Lyons 2001). This paper provides solid evidence of their permanence. The second debate is whether imperfect substitutability across different-currency assets is important empirically (see, e.g., Hau and Rey 2002). The permanent portfolio effects on price that we find here provide direct evidence of its importance. Turning to asset pricing, there has been something of a sea change in the last fifteen years regarding whether aggregate equity returns are driven primarily by shocks to fundamentals (i.e., cash flows), which is the traditional view, or by shocks to discount rates, the newer view (Campbell and Shiller 1988, Campbell 1991, Cochrane 1991; see also Vuolteenaho 2002). At the same time, people have not yet begun to consider whether all this discount-rate news might in fact originate as dispersed information that needs to be aggregated in the trading process. Our results suggest that it is time to consider this view, certainly with respect to FX, and probably with respect to equities as well.

From a modeling perspective, two main features distinguish our paper from earlier work. The first is the question we address, which is not posed clearly elsewhere in the literature, to our knowledge (whether information-based trade can still arise in a market with symmetric information about fundamentals). That the question warrants posing is evidenced by recent analysis of bond markets (e.g., Massa and Simonov 2003 and Green 2002) that uses the analysis below as its driving theoretical motivation. We extend the Lyons (1997) model to address the forecasting of discount factors (not an issue addressed in that paper). Our analysis demonstrates that inventory information is not special, but relies instead on two rather generally present ingredients: less than perfect transparency and risk aversion. Less than perfect transparency insures that inventory information is not instantly symmetrized. Risk aversion insures that inventory information is relevant to prices. The combination allows for information-based trading even when fundamental information is symmetric.

The second feature that distinguishes our modeling is our information structure. There is but a single source of information-based trade in the model: marketmakers' private knowledge of their own inventories, which are, by construction, orthogonal to fundamentals. This distinguishes our work from all three of the leading theoretical approaches to securities trading: the rational expectations approach, the microstructure information approach, and the microstructure inventory approach. Our model differs from rational expectations models (e.g., Diamond and Verrecchia

---

<sup>5</sup> Permanent price impact from inventory information would show up (and be interpreted) as fundamental information under structural approaches as well (like that of Madhavan and Smidt 1991).

1981 and Hellwig 1980) in that if one were to remove private cash-flow information from those models, aggregate endowment shocks would be fully revealed, leaving no scope for trading based on superior endowment information.<sup>6</sup> This same property is true of the recent generation of microstructure information models (e.g., Gennotte and Leland 1990, Roell 1990, Fishman and Longstaff 1992, Kumar and Seppi 1992, Madrigal 1996, and Naik, Neuberger, and Viswanathan 1999). Though designed to address quite different issues (e.g., crashes, dual trading, and transparency), all these models include private information about fundamentals as an essential ingredient: in each, superior information about some feature of the trading environment (such as risky-asset endowments or liquidity-motivated demand) allows one to forecast how price is misinterpreting cash flows. If these models excluded private information about cash flows, then knowledge of the trading environment that is unrelated to fundamentals would not be useful in forecasting price. (These models assume risk neutrality, so that direct effects on price from non-fundamental information cannot arise.) Our model departs from the microstructure inventory approach in that those models generally do not allow, or do not address, how agents would exploit their evolving inventory information via speculative position-taking (e.g., Biais 1993, Vayanos 1999, Viswanathan and Wang 2000).<sup>7</sup>

The rest of the paper is in five sections. Section I presents a trading model designed to isolate inventory information. Sections II and III describe the model's equilibrium and present results on how inventory information motivates speculative trade. Section IV presents our empirical analysis. Section V concludes.

## **I. A Model of Interdealer Trading**

Our multiple-dealer model corresponds most closely to trading in the largest multiple-dealer market—the market for spot foreign exchange (FX).<sup>8</sup> This market is an interesting target because, as noted, superior information about fundamentals—e.g., future interest differentials—is generally absent, despite market participants being convinced that superior information exists

---

<sup>6</sup> Our analysis is not the first to use endowment-type information as conditioning information (early examples include Diamond and Verrecchia 1981, Hellwig 1980, and Spiegel and Subrahmanyam 1992). The difference, as noted, is whether this constitutes superior information for speculative trade when fundamental information is symmetric (and it does not in all these other models). Kraus and Smith (1989) show that prices can move even though it is common knowledge that no new fundamental information has arrived. But they do not address the implications of non-fundamental information that is private.

<sup>7</sup> Though there is no information asymmetry across dealers in Vayanos (1998), that model does include a large non-dealer participant who trades strategically based on his endowment (which is cash flow unrelated; see also Vitale 2000). There is now a large literature on interdealer trade. On the theory side, see for example Ho and Stoll (1983), Vogler (1997), and Werner (1997).

<sup>8</sup> The model should not be viewed as applying to this market only, however. Most bond markets and many equity markets share a similar multiple-dealer structure.

(see, e.g., Cheung and Wong 2000) and despite empirical findings consistent with information asymmetries.<sup>9</sup>

We consider an infinitely lived, discrete-time trading environment with three agent types:  $n$  dealers (who behave strategically), a continuum of liquidity-demanding customers, and a continuum of liquidity-supplying customers. All agents maximize the following utility function:

$$(1) \quad U_t = \sum_{\tau=t}^{\infty} -\delta^{\tau-t} \exp(-\gamma c_{\tau})$$

where  $c_t$  is the agent's consumption in period  $t$ ,  $\gamma$  is the coefficient of (absolute) risk aversion, and  $\delta$  is the rate of time preference (the latter two parameters assumed equal for all agents). We assume that  $\delta(1+r) = 1$ , where  $r$  is the net return on the riskless asset (this assumption greatly simplifies the algebra in the dynamic programming problem).

Each day includes four trading rounds (see Figure 1). The day opens with a round of trading between dealers and liquidity-demanding customers, which is followed with a round of interdealer trading. After the first round of interdealer trading, order flow information is observed and there is a second round of interdealer trading. After the two interdealer trading rounds, there is a final round of customer-dealer trade (this time with liquidity-supplying customers) to share risk with the public more broadly. Indeed, the dealers will end each day with no net risky position, leading them to maximize expected utility over intraday returns. Daily closing prices change due to the permanent inventory shifts of the liquidity-demanding customers on the one hand and to the willingness of the liquidity-supplying customers to absorb them on the other.

A key feature of the model is that trading within a round occurs simultaneously. This simultaneous-trade approach is in the spirit of simultaneous-move games (cf, sequential-move games). Simultaneous trading has the effect of constraining dealers' conditioning information: within any round dealers cannot condition on that round's realization of others' trades. We consider this level of conditioning information more realistic than that implicit in rational-expectations models (see Hellwig 1982 for another method of relaxing the strong assumption about conditioning information in rational-expectations models). Realism aside, though, the essential implication is that constraining conditioning information in this way allows dealers to

---

<sup>9</sup> Though there is strong evidence that information asymmetries exist, these empirical results are not specific regarding their nature. See, e.g., Lyons (1995), Yao (1998), Ito et al. (1998), Covrig and Melvin (2002), Payne (1999), Evans (2002), and Evans and Lyons (2002).

trade on inventory information *before it is reflected in price*. Thus, unlike traditional trading models, in our model dealers can exploit inventory information in their trading.<sup>10</sup>

There are two assets, one riskless and one risky. The daily interest rate (net) on the riskless asset is  $r$ .<sup>11</sup> The cash flow in period  $t$ ,  $R_t$ , on the risky asset is realized prior to the last round of trading between the public and the dealers, where  $R_t$  is normally distributed about 0 with known variance  $\sigma_R^2$ :

$$(2) \quad R_t \sim N(0, \sigma_R^2)$$

The cash flows on the risky asset are serially uncorrelated and the risky asset is in zero supply initially.<sup>12</sup> Each day there are ten events, which occur in the following sequence (see Figure 1):

Round 1

- Dealers quote
- Customers trade with dealers

Round 2

- Dealers quote
- Dealers trade with dealers
- Interdealer order flow is observed

Round 3

- Dealers quote
- Dealers trade with dealers
- The cash flow  $R_t$  is realized

Round 4

- Dealers quote
- Dealers trade again with the public to share risk

Quoting Rules

The first event in all rounds is dealer quoting. Let  $P_{ij}$  denote the quote of dealer  $i$  in round  $j$  on day  $t$ . The rules governing dealer quotes are:

- (R1) Quoting is simultaneous, independent, and required
- (R2) Quotes are observable and available to all participants
- (R3) Each quote is a single price at which the dealer agrees buy and sell any amount<sup>13</sup>

---

<sup>10</sup> The lack of an opportunity to exploit inventory information in traditional microstructure models may help explain why that literature has paid so little attention to it for so long. Less than complete transparency is another means of allowing dealers to trade on inventory information before it is impounded in price.

<sup>11</sup> We assume that interest accrues between trading rounds 3 and 4. This does not affect any of our results and is assumed for simplicity of exposition.

<sup>12</sup> Adding dynamics to equation (2) would not change the essential economics of our analysis.

<sup>13</sup> In the FX market, this assumption of a single price is much less extreme than in other markets: bid-ask spreads are typically only 1-2 basis points for \$10 million trades. In any event, our basic results go through in a variation of the model in which each interdealer quote is a price schedule, but is observable and available to only one other dealer. Though this variation is quite near to reality, it is a good deal more cumbersome.

Rule 1 places this model in the simultaneous-trade approach to dealer markets (c.f., the sequential trade approach of Glosten and Milgrom 1985). Though the sequential-trade approach is popular for single-dealer modeling, it becomes unwieldy in multiple-dealer settings. (As an empirical matter, simultaneous moves in the foreign exchange and bond markets occur, for example, via electronic dealing systems that allow simultaneous quotes and simultaneous trades.) The key implication of rule R1 is that the quote of dealer  $i$  in a given round  $j$  cannot be conditioned on the round- $j$  quotes of other dealers. That quotes are required prevents a dealer from exiting the game at times of informational disadvantage; it squares with the fact that in actual multiple-dealer markets, refusing to quote violates an implicit contract of reciprocal immediacy and can be punished (e.g., by reciprocating with refusals).

Rule R2 defines the first of the model's three dimensions of transparency: quotes are fully transparent, which includes quotes for customer trades and for interdealer trades. Regarding rule R3, note that it would be straightforward to add a commission or size-dependent spread for customer trades in order to model the entry decision and endogenize the number of participating dealers. Given the focus of this paper, modeling the entry decision is a distraction so we abstract from it here.

### Customer Trades

Market orders from liquidity-demanding (LD) customers arrive in round one and are independent of the daily cash flow  $R_t$ . They are cleared at the receiving dealer's round-one quote  $P_{it}$ . Each customer trade is assigned—or preferenced—to a single dealer (possibly due to unmodeled bilateral relationships). The total (net) customer order received by a particular dealer is distributed normally around 0, with known variance  $\sigma_x^2$ :

$$(3) \quad x_{it} \sim N(0, \sigma_x^2)$$

We use the convention that  $x_{it}$  is positive for net customer purchases and negative for net sales. Though the assumptions of preferencing and exogeneity of  $x_{it}$  appear strong at this stage, we show below that equilibrium dealer quotes in round one are all the same, and conditional on public information this price is also unbiased, so LD customers are not being exploited at these prices.

LD customer trades embody the second of the model's three dimensions of transparency:  $x_{it}$  is not observed by other dealers. This is important—these customer trades are the private non-



fundamental information in the model. In foreign exchange, dealers have no direct information about other banks' customer trades.<sup>14</sup>

### Interdealer Trading

The model's four-round structure is designed to highlight the two rounds of interdealer trading that occur in the middle of each day (as this is where price discovery occurs). Let  $T_{ijt}$  denote the net outgoing interdealer order placed by dealer  $i$  in round  $j$  ( $j=2,3$ ); let  $T'_{ijt}$  denote the net incoming interdealer order received by dealer  $i$  in round  $j$ , placed by other dealers. The rules governing interdealer trading for both rounds are as follows:

- (R4) Trading is simultaneous and independent
- (R5) Trading with multiple partners is feasible
- (R6) Trades are divided equally among dealers with the same quote (if any transaction is desired at this quote).

Rule R4 generates an interesting role for  $T'_{ijt}$  in the model: because interdealer trading is simultaneous and independent,  $T_{ijt}$  is not conditioned on  $T'_{ijt}$ , so  $T'_{ijt}$  is an unavoidable disturbance to dealer  $i$ 's position in round  $j$  that must be carried into the following round. The restriction in rule R6—that trades are split equally if quotes are common—can be relaxed. For example, allowing a split into  $k < n$  equal fractions is straightforward as long as  $k$  is known. (An unknown  $m$  generates a non-normal position disturbance.) This relaxation would provide less risk sharing, but would not affect the path of price, nor would it affect dealers' incentive to speculate on the basis of private non-fundamental information (as we will see below).

Consider now the determination of dealer  $i$ 's outgoing interdealer orders in rounds 2 and 3 of each day. We sign all orders according to the party initiating the trade. Thus,  $T_{ijt}$  is positive for dealer  $i$  purchases, and  $T'_{ijt}$  is positive for purchases by other dealers from dealer  $i$ . Consequently, a positive  $x_{it}$  or  $T'_{ijt}$  corresponds to a dealer  $i$  sale. Letting  $D_{ijt}$  denote dealer  $i$ 's speculative demand, we have by definition:

$$(4) \quad T_{i2t} = D_{i2t} + x_{it} + E[T'_{i2t} | \Omega_{i2t}]$$

$$(5) \quad T_{i3t} = (D_{i3t} - D_{i2t}) + (T'_{i2t} - E[T'_{i2t} | \Omega_{i2t}]) + E[T'_{i3t} | \Omega_{i3t}]$$

where  $\Omega_{i2t}$  and  $\Omega_{i3t}$  denote dealer  $i$ 's information sets at the time of trading in round two and three, respectively. Equation (4) clarifies that dealer orders include both a speculative component  $D_{i2t}$  and two inventory control components  $x_{it}$  and  $E[T'_{i2t} | \Omega_{i2t}]$ .  $D_{i2t}$ , the speculative demand, is purely information driven: dealers use information from their private observation of customer

---

<sup>14</sup> Our specification accords well with practitioner survey responses that "better information" and "large customer base" are the two main sources of competitive advantage for large players in the FX market (Cheung and Wong 2000).

demand  $x_{it}$  to forecast subsequent price changes. The first inventory control component,  $x_{it}$ , must be fully offset in interdealer trading to establish the desired position  $D_{i2t}$ . The second inventory control component  $E[T'_{i2t} | \Omega_{i2t}]$  reflects that dealers also do their best to hedge against the incoming dealer order  $T'_{i2t}$  (the realization of which they cannot know ex-ante due to simultaneous trading). In round three, establishing the speculative demand  $D_{i3t}$  implies that the trade rule for  $T_{i3t}$  has three components. It must account for the change in speculative demand ( $D_{i3t} - D_{i2t}$ ) plus two inventory control components, the first reflecting the unexpected component of the realized incoming round-two order and the second reflecting the hedge of the round-three incoming order (the plus sign preceding  $T'_{i2t}$  in equation 5 reflects that  $T'_{i2t} > 0$  is a dealer  $i$  sale in round two).

### Interdealer Order Flow

The last event in round two defines the third of the model's three dimensions of transparency – that applying to interdealer trades. At the close of round two all dealers observe round-two interdealer order flow:

$$(6) \quad Z_t \equiv \sum_{i=1}^N T_{i2t}.$$

This sum over  $T_{i2t}$  is net interdealer demand – the difference in buy and sell orders. In bond and FX markets,  $Z_t$  is the information on interdealer order-flow provided by interdealer brokers (see Lyons 2001 for details). Actual dealers in these markets describe this as an essential source of real-time information.

Note that we specify this as a noiseless measure, which maximizes the transparency difference across trade types (customer-dealer with zero transparency and interdealer with complete transparency). As noted above, FX trades between customers and dealers do indeed have zero transparency. It is not the case, however, that the actual transparency of interdealer trades is complete. Nevertheless, it will be clear from our results that adding noise to equation (6) has no qualitative impact, so we stick to this simpler specification.

### Trading with the Public

Liquidity-supplying (LS) customer trades in round four are non-stochastic: these agents are induced to trade conditional on past price movements and other information available to them. These trades are the means by which dealers share overnight risk with the non-dealer public. Round four begins with each dealer quoting (simultaneously and independently) a scalar price  $P_{i4t}$  at which he agrees to buy and sell any amount. These quotes are observable and available to all LS customers.

The mass of public investors on the interval  $[0, m]$  is large relative to the  $n$  dealers. This implies that the dealers' capacity for bearing overnight risk is small relative to the public's capacity (despite dealers and public investors having the same coefficient of risk aversion). Nevertheless, the public is assumed to have finite risk-bearing capacity. Under these assumptions, dealers will optimally set prices such that the public willingly absorbs dealer inventory imbalances, and each dealer ends the day with no net position.<sup>15</sup> These round-four prices are conditioned on the interdealer order flow  $Z_t$ . The interdealer order flow informs dealers of the total size of the position that the public needs to absorb to bring the dealers back to a position of zero.

### Dealer Objectives and Information Sets

Each dealer determines quotes and speculative demand by maximizing the utility function shown in equation (1). Because dealers hold no overnight positions, and because dealers face a daily stochastic environment that is time invariant, solution to the corresponding one-period problem is intertemporally optimal. Letting  $W_{i4t}$  denote the end-of-day  $t$  wealth of dealer  $i$ , we can write this problem as:

$$(7) \quad \begin{aligned} & \text{MAX} \quad E[-\hat{\gamma}W_{i4t} \mid \Omega_{i1t}] \\ & \{P_{i1t}, P_{i2t}, P_{i3t}, P_{i4t}, D_{i2t}, D_{i3t}\} \\ & \text{s.t.} \end{aligned}$$

$$W_{i4t} = (W_{i0t} + x_{it}P_{i1t} + T'_{i2t}P_{i2t} + T'_{i3t}P_{i3t} - T_{i2t}P_{i2t} - T_{i3t}P_{i3t})(1+r) + (-x_{it} - T'_{i2t} + T_{i2t} - T'_{i3t} + T_{i3t})(P'_{i4t} + R_t)$$

where  $\hat{\gamma}$  is the intertemporal risk aversion coefficient (i.e., that which applies to wealth, as opposed to the  $\gamma$  in equation 1 that applies to periodic consumption).<sup>16</sup>  $P_{ijt}$  is dealer  $i$ 's round- $j$  quote, a  $'$  denotes a quote or trade received by dealer  $i$ . The first term in brackets in the constraint reflects the aggregate position in the risk-free asset at the time of the net cash flow on that asset (that cash flow being  $r$ ). The final term in brackets reflects the round-four liquidation of any remaining risky-asset position. The conditioning information  $\Omega_{ijt}$  at each decision node is summarized in Appendix A.

<sup>15</sup> Technically, these assumptions would drive dealers' overnight demand arbitrarily close to zero, but not to zero, since dealers' aggregate mass is not zero. That we treat their overnight demand as exactly zero must therefore be considered an assumption.

<sup>16</sup> The appendix shows that  $\hat{\gamma} = \gamma r(1+r)^{-1}$ . Intuitively, consumption effects from changes in wealth are smoothed across a perpetuity of consumption, which is why  $\hat{\gamma} < \gamma$ . In continuous time,  $\hat{\gamma} = r\gamma$ .

### Market Clearing

Since dealers end the day without any net position, market clearing requires that the LS customers' demand  $d_t$  absorbs all the orders of the customers in the first round. Therefore,

$$(8) \quad d_t + x_t = 0$$

where we have defined  $x_t = \sum_i x_{it}$  as the aggregate LD customer flow in round one.

To solve the model, we first conjecture the pricing and trading strategies for dealers and public investors and then show that the proposed equilibrium is consistent with dealers' and public investors' maximizing behavior. (Detailed proofs are provided in Appendix A.) Let  $h_{t-1}$  denote the aggregate holdings of the LS customers at the end of day  $t-1$ . Remember that, by assumption, dealers hold zero inventory at the end of each trading day. Thus we must have:

$$(9) \quad h_t = h_{t-1} + d_t$$

The LS customers trade dynamically in the fourth round of each day to maximize their expected utility (equation 1).

## **II. Equilibrium**

The equilibrium concept we use is that of Perfect Bayesian Equilibrium (PBE). Under PBE, Bayes rule is used to update beliefs and strategies are sequentially rational given those beliefs.

Solving for the symmetric PBE, first we consider properties of optimal quoting strategies. Consider first the sequence of end-of-day prices:

PROPOSITION 1: There exists an equilibrium in which end-of-day price is a linear function of the aggregate LS customer holdings:

$$P_{4t} = -ah_t$$

Proofs are in the appendix. The intuition is the following: Because the aggregate speculative demand of LS customers is not perfectly elastic, and given the unchanging stream of expected risky-asset cash flows ( $R_t$ ), LS customers will only hold a larger position (higher  $h_t$ ) if they are induced to do so with a lower price, i.e., with a higher expected return.

This first proposition is crucial for understanding how inventory information (i.e., non-fundamental information that is first impounded in dealer inventories) can be useful for forecasting permanent price components. LS customers will *forever* need a lower price to induce them to hold larger positions  $h_t$  indefinitely (and these positions do need to be held indefinitely

because LD customer flows are not mean reverting). Information about increments to  $h_t$  is impounded in dealer inventories and revealed to the market by via interdealer trading.

Consider now the three price quotes within each trading day. The following proposition addresses quotes in rounds one and two:

PROPOSITION 2: A quoting strategy is consistent with symmetric PBE only if the round-one and round-two quotes are common across dealers and both are unbiased predictors of the following round's price conditional on public information:

$$P_{1t} = E[P_{2t} | \Omega_{1t}]$$

$$P_{2t} = E[P_{3t} | \Omega_{2t}]$$

To understand why the quote  $P_{1t}$  to customers in round-one is an unbiased predictor of the interdealer quote  $P_{2t}$  in round two, let us suppose it is not, which we will see is inconsistent with market clearing.<sup>17</sup> Market clearing in round one requires a price at which expected dealer demands and expected customer demands sum to zero. Now, if the price change expected from round one to round two is not zero (i.e., if  $P_{1t}$  is a biased predictor of  $P_{2t}$ ), then dealers' aggregate demand in round one would not be zero. But because customer demand  $x_t$  is an exogenous, mean-zero variable, before its realization the expectation of  $x_t$  is necessarily zero. Therefore, any biased price cannot be an equilibrium price. The logic for understanding why the price in round two needs to be an unbiased predictor of the round-three price is the same: in that case, unbiasedness is necessary for market clearing in interdealer trading.

An implication of common quotes is that in round two each dealer receives a  $1/n$  share of the orders submitted by other dealers (per trading rule R6). This order corresponds to the position disturbance  $T'_{i2t}$  in the dealer's problem in Eq. (7). The next proposition addresses round-three quotes:

PROPOSITION 3: A quoting strategy is consistent with symmetric PBE only if the common round-three quote is:

$$P_{3t} = E[P_{4t}/(1+r) | \Omega_{3t}] + \lambda Z_t$$

---

<sup>17</sup> That the Nash equilibrium in price quotes is based on public information only (i.e., dealers optimally put zero weight on their private information in setting quotes) is shown in a similar setting in Lyons (1997).

The value of the constant  $\lambda$  ( $\lambda > 0$ ) is presented in the appendix. The no-arbitrage argument that establishes common quotes is the same as that for proposition 1. Like  $P_{1t}$  and  $P_{2t}$ ,  $P_{3t}$  necessarily depends only on public information (because quotes must be common). Here, the additional public information is the interdealer order-flow  $Z_t$ .

Intuition for  $\lambda > 0$  is important because  $P_{3t}$  is what motivates dealers' information-based speculation. The market clearing condition is similar to that in round two:

$$(10) \quad \sum_{i=1}^n (T_{i3t} - T'_{i3t}) = \sum_{i=1}^n (D_{i3t} + x_{it} + E[T'_{i2t} | \Omega_{i3t}] - T'_{i3t}) = 0$$

Taking the expectation using public information in round three (price being set on the basis of this expectation), we get:

$$(11) \quad \sum_{i=1}^n (E[x_{it} | \Omega_{3t}] + E[D_{i3t}(P_{3t}) | \Omega_{3t}]) = 0$$

In this case, however,  $E[x_{it} | \Omega_{3t}] \neq 0$  since interdealer order flow  $Z_t$  is contained in  $\Omega_{3t}$  and provides information about the realized  $x_{it}$ . A negative  $Z_t$ , for example, means that average  $T_{i2t}$  is negative – dealers are selling in interdealer trading. This implies that, prior to interdealer trading, customers sold on average; i.e., the negative  $Z_t$  reflects dealers' laying off long positions they acquired from customers – the  $x_{it}$  term in Eq. (11). If customers sold in round one, then in round two dealers are long. To clear the market in round three, expected  $P_{4t} - P_{3t}(1+r)$  must be positive to induce dealers to hold this long position.  $P_{3t}$  must therefore fall below  $E[P_{4t}/(1+r) | \Omega_{3t}]$  to provide the positive intraday return needed to compensate dealers.<sup>18</sup> The end result is that the negative  $Z_t$  drives a reduction in price, that is,  $\lambda > 0$ , as we set out to show. As in proposition 1, any price other than  $P_{3t} = E[P_{4t}/(1+r) | \Omega_{3t}] + \lambda Z_t$  is incompatible with equilibrium since  $dD_{i3t}/dP_{3t} < 0$ . In the fourth round, dealers trade with the public to clear their inventory.

Notice that the price change between the second and third rounds of trade will be positively correlated with inter-dealer order flows  $Z_t$ . Moreover, the price change between the third and fourth rounds will be negatively correlated with inter-dealer order flows  $Z_t$ , i.e., price exhibits high frequency mean reversion. If the round-one customer flow had been publicly observed, then dynamics would be unrelated to the interdealer flows.

---

<sup>18</sup> Recall that the day's random payoff  $R_t$  (equation 2) is realized between the determination of  $P_{3t}$  and  $P_{4t}$ , so dealers are exposed to intraday risk here. At the same time, the expectation of  $R_t$  is always zero (i.e., there is no private information about fundamentals), which is why we do not include  $R_t$  in this discussion of the dealers' expected return.

## Trading Strategies

Given the quoting strategy described in propositions 1 and 2, the following optimal trading strategy corresponds to symmetric linear equilibrium:

PROPOSITION 4: The following trading strategy profile  $\forall i \in \{1, \dots, n\}$  is a perfect Bayesian equilibrium, where  $\beta_2 > 1$  and  $\beta_3 < 0$ :

$$T_{i2t} = \beta_2 x_{it}$$

$$T_{i3t} = \beta_3 x_{it}$$

The values of the  $\beta$  coefficients are presented in the appendix.<sup>19</sup> Recall that the quoting rules for  $P_{2t}$  and  $P_{3t}$  are linear in  $E[P_{3t}]$ , and in  $E[P_{4t}]$  and  $Z_t$ , respectively. The trading rules have a corresponding linear structure deriving from exponential utility and normality, which generate linear speculative demand. The derivations in the appendix insure that the quoting and trading strategies are mutually consistent and sequentially rational. These strategies impound dealer recognition that their own actions affect price.

Here is the intuition behind the result. Recall from propositions 2 and 3 that the return  $P_{3t} - P_{2t}$  is wholly a function of public information revealed in that time interval. Recall too that, within that interval, interdealer order flow  $Z_t$  is the only public information revealed. The positive value of the coefficient  $\lambda$  in proposition 3 ensures that the return  $P_{3t} - P_{2t}$  is positively related to the realization of  $Z_t$ . At the time of trading in round two, individual dealers do not know the realization of  $Z_t$ . But since  $Z_t$  is simply the sum over all the dealers' round-two trades  $T_{i2t}$ , each dealer does know one component of  $Z_t$ , namely, his own trade (and dealers are fully strategic about their own effect on prices). This creates a risk-return tradeoff that would not otherwise exist, the risk being uncertainty about other dealers' trades in  $Z_t$  (and there is nothing in the model that helps a dealer forecast those other components of  $Z_t$ ). To see why  $\beta_2$  is larger than 1, i.e., why the dealer rides herd on his customer order, notice that the positive  $\lambda$  driving  $P_{3t}$  means that price will move in the same direction as the dealer's order  $T_{i2t}$ . After receiving a positive customer trade  $x_{it}$ , the dealer is short. A non-speculative dealer would buy just enough in the interdealer market to cover that short. A speculative dealer buys more than this because the long position will profit from the price increase his trade induces. Moreover,  $\beta_3 < 0$ , as dealers adjust their round-

---

<sup>19</sup> The result that  $\beta_2 > 1$  is reminiscent of results from the dual trading literature (see, e.g., Roell 1990 and Fishman and Longstaff 1992). The dual trading literature is set in a context of risk-neutral agents so it is not equipped to address non-fundamental information. That a setting without any private information about fundamentals can produce trading behavior similar to settings that include such information is part of what we set out to show in this paper.

two positions in the third round (per equation 5). This indicates that price ( $P_{3t}-P_{2t}$ ) will be positively correlated with contemporaneous order flows ( $T_{2t}$ ) but negatively correlated with lagged order flows ( $T_{1t}$ ).

### III. Speculation Based on Inventory Information

The trading strategies in proposition 4 have implications for the role of private non-fundamental information and, more specifically, for the role of inventory information. For example, the coefficient in the rule for round-two trading implies that non-fundamental information motivates dealer speculation.

PROPOSITION 5: Under some regularity conditions, the proposed equilibrium characterized in propositions 1-4 exists. In addition, though the model contains no private fundamental information, dealers can still speculate based on private non-fundamental information.

This proposition follows directly from the expression for  $\beta_2$  in proposition 4. Specifically, the appendix shows that  $\beta_2 > 1$  (under some regularity conditions). In the “no speculation” case,  $\beta_2 = 1$  since a non-speculative dealer would simply offset his round-one customer trade with an interdealer trade, one-for-one. Instead, the dealer chooses to open a risky position, using his private information as the basis.

The next proposition generalizes our simple example.

PROPOSITION 6: Superior information about other determinants of  $P_{3t}$  also qualifies as private non-fundamental information, in particular, superior information about  $n$ ,  $\gamma$ ,  $\sigma_x^2$ , or  $\sigma_R^2$ .

This proposition follows directly from the proof of proposition 2, which shows that  $P_{2t}$  is a function of all these parameters, none of which provides information about cash flows. Though our model focuses on inventory information as one element within the larger non-fundamental class, proposition 6 clarifies that many other partitions of the non-fundamental class exist (indeed, there are many partitions beyond those specified in the proposition).

The reason we specify the model without private fundamental information is to establish that our results are not based on trades that are cash-flow irrelevant being mistaken for cash-flow



relevant trades, as occurs in many other models in the literature. Given this specification choice, it is helpful to link our result to more familiar contexts, in particular, those with private fundamental information. The next proposition makes this link by addressing the robustness of private non-fundamental information as a motive for speculation.

**PROPOSITION 7:** Introducing private fundamental information does not preclude dealers from speculating based on private non-fundamental information.

This proposition also follows directly from the earlier propositions. Specifically, introducing private information about  $R_t$  does not alter the two risk-premium effects on  $P_{3t}$  necessary to clear the interdealer market (the first being the transitory component that clears the interdealer market and the second being the permanent component that anticipates the  $P_{4t}$  that clears the wider market). Clearly, introducing private fundamental information will alter the demand function described following appendix equation (A7), but one component of that demand function will still represent non-fundamental information.

Before turning to the empirical model of the next section, let us review some key features of our analytical findings. First, when customer order flows (and dealer inventories) are not publicly known, interdealer order flow is the means by which information impounded in dealer inventories is subsequently impounded in price. Second, one of the price components that inventory information produces is a transitory, intraday effect (corresponding to  $P_{4t}-P_{3t}$ ). Third, the other price component that inventory information produces is permanent, corresponding to  $P_{4t}-P_{4t-1}$  (i.e., the daily return). This sharp distinction between transitory and permanent components contrasts with analysis of finite-horizon models (e.g., Lyons 1997 and Evans and Lyons 2002) where the meaning of permanent is unclear since price ultimately converges to the ultimate payoff value in these models.

#### **IV. Empirical Model and Results**

Our empirical strategy is the following. First, we choose a financial market—foreign exchange—for which both theory and empirics line up strongly against the presence of private information of the fundamental type (as described in the introduction; for this statement we have in mind the major currencies like the \$/DM or \$/Yen). Second, we provide a flexible empirical model for decomposing price into the components highlighted in previous sections. Third, we design the empirical model so that price discovery is driven by the same order-flow concept that drives our analytical work, namely, interdealer order flow  $Z_t$ .

We begin by breaking transaction prices into three parts:

$$(12) \quad P_t = P_t^{PI} + P_t^{II} + w_t,$$

The observed price level includes a component of public fundamental information (PI),  $P_t^{PI}$ , a component of inventory information (II),  $P_t^{II}$ , and a serially uncorrelated sampling error,  $w_t$ .<sup>20</sup> (Though  $E_t[R_{t+1}]$  is by design not changing in our theoretical model, an empirical model needs to accommodate such changes, hence the introduction of  $P_t^{PI}$ .) The public information component follows a random walk:

$$(13) \quad \Delta P_t^{PI} = \varepsilon_t$$

The inventory information component  $P_t^{II}$  is impounded in price through interdealer trading; this is where dispersed information about customer demands  $x_{it}$  is aggregated in observable order flow. We model the inventory information component as following the (quite flexible) moving-average process:

$$(14) \quad \Delta P_t^{II} = \Gamma(L)e_t.$$

where  $e_t$  is the innovation in aggregate interdealer flow  $Z_t$  during period  $t$  and  $\Gamma(L)$  is a lag polynomial (as yet unspecified). This innovation  $e_t$  can be estimated from realized interdealer flow  $Z_t$  as a moving-average process:

$$(15) \quad Z_t = \Theta(L)e_t.$$

---

<sup>20</sup> This error is present because at any given moment our dataset provides a single price from what is in reality a cross-sectional distribution of prices (see Evans 2002 for analysis of this cross-sectional property of these data). One cause of this distribution is idiosyncratic effects of inventory on individual dealer prices (described in our introduction as the first of the three price effects from non-fundamental trades). These transitory, idiosyncratic effects do not arise in our theoretical model, nor are they the focus here. Our focus is instead on the price effects from inventory information that are common to all dealers (both transitory and permanent effects).

where  $\Theta(L)$  is another lag polynomial (as yet unspecified). Per our model above, the  $e_t$  shocks are reflections of (unmeasured) innovations in aggregate LD customer orders  $x_t$ . Importantly, these shocks are independent of the public information shocks  $\varepsilon_t$  and the sampling errors  $w_t$ .

To disentangle transitory and permanent effects of inventory information, we decompose that component of price change  $P_t''$  into two parts, one transitory and one permanent. In particular, we assume that (14) can be written as:

$$(16) \quad \Delta P_t'' = \bar{\Gamma}(L)e_t + \Gamma(1)e_t$$

where the lag polynomial  $\bar{\Gamma}(L)$  is defined as  $\bar{\Gamma}(L) \equiv \Gamma(L) - \Gamma(1)$ . The notation  $\Gamma(1)$  is standard for denoting the sum of all coefficients in a moving-average polynomial  $\Gamma(L)$ . As such,  $\Gamma(1)$  identifies the permanent change in  $P_t''$  induced by an  $e_t$  shock (i.e., once  $e_t$  has passed through all of the moving-average coefficients, the final net effect is the sum of all those coefficients). Thus, the first effect of inventory information  $\bar{\Gamma}(L)e_t$  is transitory and the second  $\Gamma(1)e_t$  is permanent. In our earlier theoretical model, these two parts correspond to the two price effects engendered by the initial impact of  $\lambda Z$ . The transitory part in that model corresponds to an intraday risk premium (in the spirit of Spiegel and Subrahmanyam 1995), whereas the permanent part is that which remains even after dealer inventories are shared marketwide (i.e.,  $P_{4t} - P_{4t-1}$ ).<sup>21</sup> Both parts are driven by order flow shocks but differ in their long-run impact on the price level. The focus of our empirical analysis will therefore be on estimating  $\Gamma(1)$  and  $\bar{\Gamma}(L)$ . If inventory information plays a role in determining prices, we should expect to find significant coefficients in both  $\Gamma(1)$  and  $\bar{\Gamma}(L)$ .<sup>22</sup>

### Estimation

To estimate equations (12) – (15) we need to restrict the form of the moving-average polynomials  $\Gamma(L)$  and  $\Theta(L)$ . We begin with the following flexible forms (and provide results for more restricted forms, which ensures robustness has not been compromised):

---

<sup>21</sup> As noted in the introduction, inventory effects on price are of three types. This empirical specification is designed to separate the permanent effects from the transitory effects. (Our theoretical model does not admit the first type because idiosyncratic price effects are ruled out by no-arbitrage.)

<sup>22</sup> One might argue that a finding of significant  $\bar{\Gamma}(L)$  is not sufficient to establish the presence of speculative opportunities of the kind described by our model. For example, if customer trades are perfectly transparent then superior information would not exist, even in a world where the estimated  $\bar{\Gamma}(L)$  is positive. But it is an institutional

$$(17) \quad \Gamma(L) = \frac{B(L)}{A(L)} = \frac{b_0 + b_1L + b_2L^2}{1 - a_1L - a_2L^2}$$

$$(18) \quad \Theta(L) = \frac{D(L)}{C(L)} = \frac{1 + d_1L + d_2L^2}{1 - c_1L - c_2L^2}$$

where the roots of  $A(z)=0$ , and  $C(z)=0$  lie outside the unit circle. Thus, the inventory information component in prices,  $\Delta P_t^H$ , and interdealer order flow  $Z_t$  are assumed to follow covariance stationary processes. Substituting these restrictions into equations (12) – (15) allows us to write the model's reduced form as:

$$(19) \quad \Delta P_t = A(L)^{-1} B(L) e_t + \varepsilon_t + w_t - w_{t-1},$$

$$(20) \quad Z_t = C(L)^{-1} D(L) e_t.$$

The parameters to be estimated include the coefficients of the lag polynomials  $A(L), B(L), C(L)$ , and  $D(L)$ , and the variances of the shocks  $w_t, \varepsilon_t$ , and  $e_t$ . Estimates of these parameters are obtained by the Generalized Method of Moments (GMM) using the autocovariance matrix for  $Y_t = [\Delta P_t \ Z_t]'$ . In particular, let  $\Omega(j) = E[Y_t Y_{t-j}']$  denote the covariance matrix of  $Y_t$  at lag  $j$ . GMM estimates are obtained by matching the theoretical values for  $\Omega(j)$  implied by (19) and (20) against sample estimates of  $\Omega(j)$  for  $j = 0, 1, \dots, 12$  (see appendix for details).

Our empirical model differs in several respects from those in earlier studies based on time-series models for price changes and order flow, most notably Hasbrouck (1991), and more recently in the context of FX by Payne (1999). Hasbrouck's calculations of the share of permanent price changes attributable to order flow are derived from estimates of a VAR. This approach assumes that the Wold (Vector Moving Average) representation for price changes and order flow is invertible. In our model, by contrast, the VMA process is not invertible due to the presence of the sampling-error component (i.e., the  $w_t$  shock) in transaction prices. (Evans, 2002

---

fact that FX dealers have private information about their customers' trades, so this critique of our findings is of little practical relevance.

shows that this component is economically quite significant, a finding confirmed by the parameter estimates in Table 1 below.) The second important difference in our approach is that it allows us to place theoretical restrictions on the VMA representation for price changes and order flow implied by theory. In particular, our model implies that there are no lagged values of  $\varepsilon_t$  in the VMA representation. It is not possible to place these restrictions on the VMA representation implied by a VAR (without parameter restrictions). Thus, even if our model had an invertible VMA representation, the VAR method would yield less efficient estimates of the price change and order flow process (under the null of a correctly specified model). Our model also differs from Evans (2002) in that we jointly estimate the price change and order flow processes. This allows us to conduct inference on the response of price to inventory shocks; and, in particular, assess the significance of any permanent price effects.

### Data

We use transactions prices and interdealer order flow from Reuters Dealing 2001 observed at a five-minute observation interval. (See Evans 2002 for more detail on the Reuters source of these data.)  $P_t$  is the last deutsche mark purchase price for dollars during observation interval  $t$ . (Using only purchase prices, i.e., only transactions at the ask, eliminates noise from bid-ask bounce.) This is the transaction price of buyer-initiated transactions between two dealers. Interdealer order flow,  $Z_t$ , is the difference between the number of buyer- and seller-initiated orders between the midpoints of intervals  $t$  and  $t+1$ .<sup>23</sup>

### Results

Table 1 reports results for 6 alternative specifications for the lag polynomials  $A(L)$ ,  $B(L)$ ,  $C(L)$ , and  $D(L)$ . Panel A reports the GMM parameter estimates and asymptotic standard errors corrected for the presence of conditional heteroskedasticity. Models I-III assume that order flow follows an ARMA(2,1) process, while models IV – VI assume an ARMA(2,2) process.<sup>24</sup> Estimates of the process  $\Delta P_t^I$  (price increments from inventory information) are reported in the first four rows of the table. The estimate of  $b_0$  is approximately 0.30 in every model and is highly statistically significant. This implies that shocks to interdealer order flow

---

<sup>23</sup> Our dataset does not include information on trade size, only trade sign. Even in instances where past researchers have had data on both sign and size, however, they have chosen to measure order flow as the net number of trades (e.g., Hasbrouck 1991).

<sup>24</sup> As shown in equation (17), we allowed for the lag polynomial  $B(L)$  to be second order, but there is no evidence that  $b_2 \neq 0$ , so our results are based on  $B(L)$  being first order.

affect prices contemporaneously. The next four rows report estimates of the interdealer order flow process  $Z_t$ . These estimates imply that  $Z_t$  is a persistent but stationary process.

Panel B reports implications for permanent and transitory price effects from inventory information. The first row captures the permanent impact. It reports the implied estimate of  $\Gamma(1)$ .<sup>25</sup> The estimates range from 0.00108 to 0.00189 and are all highly statistically significant. Given an average trade size of \$3.9 million, these estimates imply that a \$1 billion positive shock to interdealer order flow permanently increases the deutsche mark price of the dollar by 0.25 to 0.45 of a pfennig. The next rows report the estimates of  $\Gamma(0)$ , which correspond to the total *initial* impact of order flow shocks (i.e., the permanent and transitory combined). These estimates range from 0.00270 to 0.00306 and are all significantly different from zero. They imply that a \$1 billion positive shock to interdealer order flow contemporaneously increases the deutsche mark price of a dollar by 0.69 to 0.78 pfennig (about 50 basis points).

For further perspective on the economic significance of shocks to inventory information, we compare the response of price to this information relative to the price response to public information. Specifically, we use the model estimates to compute  $IR(k)$ : the response of the price level at  $t+k$  to a one standard deviation shock in inventory information  $e_t$ , relative to the effects of a one standard deviation shock in public information  $\varepsilon_t$ . As the table shows, the immediate impact of a typical inventory shock is approximately 45 percent of the size of a typical public news shock. Over the next 30 minutes, the transitory contribution of the inventory shock subsides. Thereafter, prices reflect the permanent effect of the inventory shock, which is between 15 and 30 percent of the size of a typical public news shock (depending on the model specification). These estimates of permanent effects indicate that inventory shocks play an economically significant role in the determination of spot rates at macroeconomic frequencies.

Note that the long-run price effects are precisely estimated and highly statistically significant.<sup>26</sup> We attribute this level of precision to the four-month span of our data. Though this time span is short relative to many macroeconomic half-lives (i.e., the half-lives of stationary macro time-series), it is orders of magnitude longer than the half-lives of variables associated

---

<sup>25</sup> See Appendix B for details on our econometric identification of the permanent and transitory components of the lag polynomial  $\Gamma(L)$ . Specifically, we examine specifications flexible enough to allow short lags (up to 15 minutes) before the permanent effects of order flow are fully reflected in prices.

<sup>26</sup> The precision of the estimates differentiates our results on permanent effects from others in the literature. In particular, Payne (1999) computes the contribution of order flow shocks to exchange rates from a VAR using a sample of one week. His data on brokered interdealer trading offer less information on permanent price effects than our four-month sample. In one section of his paper, Evans (2002) reports variance decompositions indicating that shocks to order flow affect prices at various frequencies, but does not address the statistical significance of permanent effects. (Froot and Ramadorai 2002 address permanent effects from portfolio flows, but their data do not include measures of order flow, i.e., there is no way to distinguish buyer- from seller-initiated trades, which is essential for disentangling demand curves shifts—news—from price-induced movement along demand curves.)

with trading.<sup>27</sup> For example, Lyons (1995) estimates the half-life of an FX dealer's positions to be approximately 10 minutes. On par with this, our estimates of the impulse response functions show that the half-life of transitory effects from inventory shocks is about 15 minutes. Compared against this trading time scale, four months of trading data constitute a rather long sample, providing further assurance that our inferences concerning persistent effects are reliable.

We examine the temporal impact of inventory shocks more closely in the remaining rows of the table. Here we report the share of the total price effect from inventory information that comes from the transitory component. For this purpose we calculate the ratio:

$$(21) \quad R(k) = \frac{\sqrt{\text{var}(\Delta^k P_t^{Trans})}}{\sqrt{\text{var}(\Delta^k P_t^I)}}$$

where  $\Delta^k$  denotes the  $k$ -difference operator and  $P_t^{Trans}$  is the transitory component  $\bar{\Gamma}(L)e_t$  (from equation 16). The table reports ratios ranging from  $k = 6$  to 24 (5 minutes to 2 hours), together with asymptotic standard errors.

If transitory effects from inventory information are not an economically important factor driving prices in the short run, the estimates of  $R(k)$  should be insignificantly different from zero for all horizons  $k$ . The table shows that this is not the case (for any of the models): the transitory effect of inventory information appears to contribute between 43 and 89 percent of the standard deviation in inventory information price changes over horizons from 30 minutes to two hours.

---

<sup>27</sup> To elaborate, consider the half-life of departures from the macro relation called purchasing power parity (PPP). The consensus estimate of that half-life is 4 years (Froot and Rogoff 1995). With such a long half-life, how large a sample would be needed to estimate PPP with as many half-lives in the sample as we have in our sample? The estimated half-life of transitory inventory effects in our sample is 15 minutes. Our four-month sample implies about 3500 of these half-lives (based on trading between 7:00 am and 5:00 pm). The comparable time span for macro PPP data would be approximately 14,000 years (3500 x 4 years). While we cannot claim that inventory shocks literally affect the exchange rate forever (since that would require infinite data series), we feel justified in asserting that our model provides precise estimates of the persistent effects.

**Table 1: Model Estimates**

	Model Variations					
	I	II	III	IV	V	VI
	Coeff. (S.E.)	Coeff. (S.E.)	Coeff. (S.E.)	Coeff. (S.E.)	Coeff. (S.E.)	Coeff. (S.E.)
<b>Panel A</b>						
$a_1$	-0.646 (0.048)	-0.529 (0.073)	-0.341 (0.346)	-0.623 (0.048)	-0.457 (0.076)	-0.374 (0.319)
$a_2$			0.123 (0.251)			0.063 (0.225)
$b_0$ (x100)	0.291 (0.035)	0.274 (0.036)	0.270 (0.036)	0.306 (0.034)	0.290 (0.036)	0.288 (0.036)
$b_1$ (x100)		-0.100 (0.037)	-0.153 (0.088)		-0.124 (0.037)	-0.147 (0.088)
$c_1$	1.039 (0.081)	1.049 (0.089)	1.057 (0.094)	0.423 (0.153)	0.424 (0.151)	0.428 (0.155)
$c_2$	-0.129 (0.035)	-0.125 (0.036)	-0.128 (0.036)	0.363 (0.107)	0.382 (0.104)	0.379 (0.106)
$d_1$	-0.850 (0.073)	-0.876 (0.081)	-0.885 (0.087)	-0.211 (0.151)	-0.226 (0.149)	-0.231 (0.152)
$d_2$				-0.442 (0.088)	-0.473 (0.089)	-0.470 (0.090)
$\sigma_e$	4.233 (0.080)	4.217 (0.080)	4.212 (0.080)	4.231 (0.081)	4.215 (0.080)	4.215 (0.081)
$\sigma_\varepsilon$	0.038 (0.002)	0.037 (0.002)	0.037 (0.002)	0.038 (0.002)	0.039 (0.002)	0.038 (0.002)
$\sigma_\omega$	0.034 (0.002)	0.033 (0.002)	0.034 (0.002)	0.034 (0.002)	0.033 (0.002)	0.033 (0.002)
<b>Panel B</b>						
$\Gamma(1)$ (x100)	0.177 (0.025)	0.114 (0.032)	0.096 (0.042)	0.189 (0.025)	0.114 (0.032)	0.108 (0.040)
$\Gamma(0)$ (x100)	0.291 (0.035)	0.274 (0.036)	0.270 (0.036)	0.306 (0.034)	0.290 (0.036)	0.288 (0.036)
$IR(0)$ : impact	0.463 (0.052)	0.436 (0.051)	0.430 (0.052)	0.482 (0.051)	0.440 (0.049)	0.448 (0.050)
$IR(3)$ : 15 min.	0.357 (0.036)	0.253 (0.051)	0.225 (0.057)	0.369 (0.036)	0.229 (0.050)	0.228 (0.056)
$IR(6)$ : 30 min.	0.261 (0.042)	0.170 (0.049)	0.141 (0.068)	0.279 (0.042)	0.168 (0.047)	0.160 (0.064)
$IR(\infty)$ : perm.	0.281 (0.037)	0.181 (0.049)	0.153 (0.065)	0.297 (0.038)	0.173 (0.047)	0.167 (0.060)
$R(6)$ : 30 min.	0.796 (0.013)	0.864 (0.038)	0.892 (0.064)	0.790 (0.012)	0.866 (0.038)	0.878 (0.058)
$R(12)$ : 1 hr.	0.591 (0.013)	0.689 (0.059)	0.736 (0.111)	0.585 (0.012)	0.698 (0.060)	0.715 (0.093)
$R(24)$ : 2 hrs.	0.430 (0.011)	0.525 (0.061)	0.576 (0.125)	0.425 (0.010)	0.535 (0.064)	0.554 (0.101)

Notes: The table reports GMM estimates of the model:

$$\Delta P_t = \Gamma(L)e_t + \varepsilon_t + w_t - w_{t-1}, \quad Z_t = \Theta(L)e_t$$

where  $\Gamma(L) = (b_0 + b_1L)/(1 - a_1L - a_2L^2)$  and  $\Theta(L) = (1 + d_1L + d_2L^2)/(1 - c_1L - c_2L^2)$ .  $P_t$  is the last DM purchase price for dollars during observation interval  $t$ . Interdealer order flow,  $Z_t$ , is defined by the difference between the number of buyer-initiated orders and seller-initiated orders between the midpoints of intervals  $t$  and  $t+1$ . In Panel B,  $\Gamma(1)$  measures the permanent price effect of shocks to inventory information, whereas  $\Gamma(0)$  measures the initial price impact of shocks to inventory information.  $IR(k)$  is the impulse response of the price level at  $t+k$  to a one standard deviation shock to inventory information  $e_t$ , relative to the effects of a one standard deviation shock to public information  $\varepsilon_t$ .  $R(k)$  is the contribution at horizon  $k$  of transitory price changes from inventory information to the standard deviation of total price changes from inventory information (measured in 5 minute intervals).

$$R(k) = \left( \sqrt{\text{var}(\Delta^k P_t^{Trans})} \right) \left( \sqrt{\text{var}(\Delta^k P_t^{II})} \right)^{-1},$$

Estimation details are described in Appendix B. Asymptotic standard errors, corrected for conditional heteroskedasticity, are reported in parentheses.



## V. Conclusion

We posed a specific question at the outset: In a market with symmetric information about fundamentals, can information-based trade still arise? This paper shows that the answer is yes and provides some conceptual room for why. That there is any need for this conceptual room is due at least in part to models of trading that historically define private information rather narrowly. The inventory information we introduce here is unrelated to fundamentals, yet it still provides a motive for speculation because it correlates with future price. As is well known, many factors affect future price not by affecting expected cash flows, but by affecting discount rates. Examples include features of the economy like traders' endowments, risk aversions, trading constraints, and many others. What we have done here is add superior knowledge of these factors. Note that our model does not rely on non-fundamental trades masquerading as fundamental information; to make this distinction clear we specified the model without any private fundamental information. In effect, the paper occupies a space between literatures in asset pricing and microstructure, the former treating permanent endowment effects on price but without information asymmetries (e.g., Duffie and Constantinides 1996 and Storesletten et al. 2003), whereas the latter treats information asymmetries but only those relating to cash flows.

The idea of inventory information is not special, but it relies instead on two rather generally present ingredients: less than perfect transparency and risk aversion. Less than perfect transparency insures that inventory information is not instantly symmetrized. Risk aversion insures that inventory information is relevant to prices. The combination allows for trading based on superior information even when fundamental information is symmetric.<sup>28</sup>

Our analysis clarifies that prices effects from non-fundamental trades are of three distinct types. The first is transitory and idiosyncratic to individual dealer prices—the so-called inventory effects from microstructure theory. The second type is transitory and common to all dealer prices (intraday risk premia). The third type is permanent and common to all dealer prices. This third effect on price persists even after dealer inventories are fully shared economy-wide (i.e., the portfolio effects that arise from imperfect substitutability across risky assets; see, e.g., Branson and Henderson 1985). These traditional models do not address the third type of effect because inventory risk in these models is assumed diversifiable at the economy-wide level (or because these models do not allow market-wide risk sharing to occur—dealers are forced to hold the risky

---

<sup>28</sup> As noted, for readers concerned about the single-price assumption used for interdealer trade, our basic results go through in a variation of the model in which each dealer quotes a price schedule to one other dealer that is not observable market-wide.

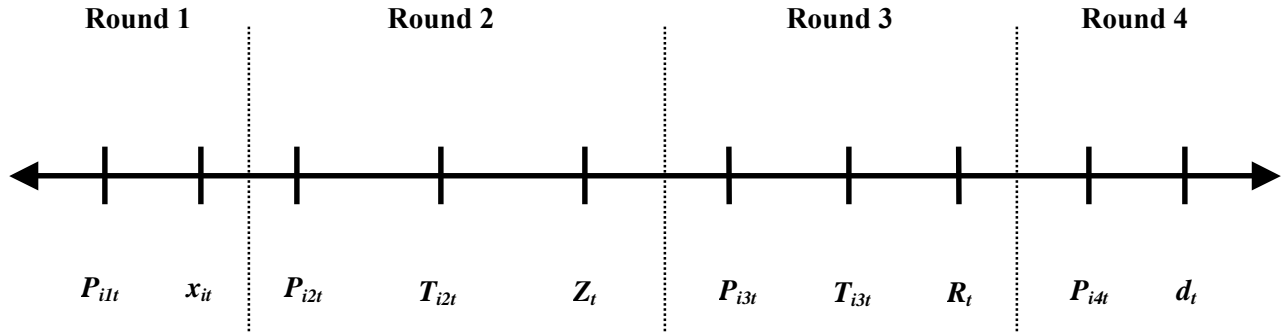
asset at the time of cash flow realization). Our empirical methodology is designed to accommodate this third, permanent effect.

We find that inventory information has both transitory and permanent price effects. For permanent effects, our estimates imply that a \$1 billion positive shock to interdealer order flow increases the DM price of a dollar in the long run by 0.25 to 0.45 of a pfennig (depending on model variations). These permanent effects also contribute significantly to the variance of permanent price changes, ranging from 15 to 30 percent. Transitory price effects from inventory information are also significant: of the total variance of price changes from inventory information, transitory effects account for between 43 and 89 percent at horizons from 30 minutes to two hours. In the end, these results suggest that past findings of private information in trades (whether in FX or other security markets) may have been picking up information that is non-fundamental, at least in part, despite the common interpretation of the information as wholly fundamental.

The results of this paper have implications in exchange rate economics and asset pricing more broadly. In exchange rate economics, the results speak directly to two big debates. The first is the degree to which order-flow effects on price persist. This paper provides solid evidence of their permanence. The second debate is whether imperfect substitutability across different-currency assets is important empirically. The permanent portfolio effects on price that we find here provide direct evidence of its importance. In asset pricing, there is an emerging consensus that aggregate equity returns are driven primarily by shocks to discount rates. At the same time, the asset pricing literature has not yet considered whether all this discount-rate news might in fact originate as dispersed information that needs to be aggregated in the trading process. Our results suggest that it is time to consider this view, certainly with respect to FX, and probably with respect to equities and bonds as well.

**Figure 1**

**Model Timing (each day)**



**Notation**

$P_{i1t}$ : dealer  $i$ 's quote in round one of day  $t$ .

$x_{it}$ : net customer order received by dealer  $i$ .

$P_{i2t}$ : dealer  $i$ 's quote in round two.

$T_{i2t}$ : dealer  $i$ 's net outgoing order to other dealers in round two.

$Z_t$ : net interdealer order flow in round two.

$P_{i3t}$ : dealer  $i$ 's quote in round three

$T_{i3t}$ : dealer  $i$ 's net outgoing order to other dealers in round three.

$R_t$ : risky asset cash flow realized.

$P_{i4t}$ : dealer  $i$ 's quote in round four.

$d_t$ : aggregate customer orders in round four (market clearing requires  $\sum_i x_{it} = d_t$ )

## References

- Andersen, T., T. Bollerslev, F. Diebold, C. Vega, 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange, *American Economic Review*, 93, 38-62.
- Biais, Bruno, 1993, Price formation and equilibrium liquidity in fragmented and centralized markets, *Journal of Finance* 48: 157-184.
- Branson, William, and Dale Henderson, 1985, The specification and influence of asset markets, in *Handbook of International Economics*, vol. 2, R. Jones and P. Kenen (eds.), North Holland: Amsterdam.
- Campbell, J., 1991, A variance decomposition for stock returns, *Economic Journal*, 101, 157-179.
- Campbell, J., and R. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1, 195-228.
- Cheung, Yin-Wong, and Clement Wong, 2000, A survey of market practitioners' views on exchange rate dynamics, *Journal of International Economics*, 51, 401-423.
- Ciccotello, C. S. and F. Hatheway, 2000, Indicating Ahead: Best Execution and the NASDAQ Preopening, *Journal of Financial Intermediation*, 9, 184-212.
- Cochrane, J., 1991, Volatility tests and efficient markets: a review essay, *Journal of Monetary Economics*, 27, 463-485.
- Covrig, V., and M. Melvin, 2002, Asymmetric information and price discovery in the FX market: Does Tokyo know more about the Yen? *Journal of Empirical Finance*, 9, 271-285.
- Diamond, D., and R. Verrecchia, 1981, Information aggregation in a noisy rational expectations economy, *Journal of Financial Economics*, 9: 221-235.
- Duffie and Constantinides, 1996, Asset Pricing with heterogeneous consumers, *Journal of Political Economy*, 104: 219-240.
- Evans, Martin, 2002, FX trading and exchange rate dynamics, *Journal of Finance*, 57: 2405-2448.
- Evans, Martin, and Richard Lyons, 2002, Order flow and exchange rate dynamics, *Journal of Political Economy*, 110: 170-180.
- Fama, Eugene, and Kenneth French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246-273.
- Fishman, M., and F. Longstaff, 1992, Dual trading in futures markets, *Journal of Finance*, 47: 643-672.
- Foster, Douglas, and S. Viswanathan, 1990, A theory of interday variations in volumes, variances, and trading costs in securities markets, *Review of Financial Studies* 3, 593-624.

- Foster, Douglas, and S. Viswanathan, 1993, Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models, *Journal of Finance* 48, 187-211.
- Frankel, Jeffrey, and Andrew Rose, 1995, Empirical research on nominal exchange rates, in G. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, Elsevier Science: Amsterdam, 1689-1729.
- French, Kenneth, and Richard Roll, 1986, Stock return variance: The arrival of information and the reaction of traders, *Journal of Financial Economics* 17, 99-117.
- Froot, Kenneth, and Tarun Ramadorai, 2002, Currency returns, institutional investor flows, and exchange rate fundamentals, NBER Working Paper 9101, August.
- Froot, Kenneth, and Kenneth Rogoff, 1995, Perspectives on PPP and Long-Run Real Exchange Rates, in G. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, Elsevier Science: Amsterdam, 1647-1688.
- Genotte, Gerard, and Hayne Leland, 1990, Market liquidity, hedges, and crashes, *American Economic Review*, 80, 999-1021.
- Glosten, Lawrence, and Lawrence Harris, 1988, Estimating the components of the bid-ask spread, *Journal of Financial Economics* 21, 123-142.
- Glosten, Lawrence, and Paul Milgrom, 1985. Bid, ask, and transaction prices in a specialist market with heterogeneously informed agents, *Journal of Financial Economics* 14, 71-100.
- Green, C., 2003, Economic news and the impact of trading on bond prices, typescript, Emory University, May, *Journal of Finance*, forthcoming.
- Grossman, S., 1976, On the efficiency of competitive stock markets when agents have diverse information, *Journal of Finance*, 31: 573-585.
- Grossman, S., and M. Miller, 1988, Liquidity and market structure, *Journal of Finance*, 43: 617-633.
- Grossman, S., and J. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review*, 70: 393-408.
- Hansch, O., N. Naik, and S. Viswanathan, 1998, Do inventories matter in dealership markets? Evidence from the London stock exchange, *Journal of Finance* 53, 1623-1656.
- Hasbrouck, Joel, 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179-208.
- Hau, H., and H. Rey, 2002, Exchange rates, equity prices, and capital flows, NBER Working Paper 9398, December.
- Hellwig, Martin, 1980, On the aggregation of information in complete markets, *Journal of Economic Theory*, 26: 279-312.

- Hellwig, Martin, 1982, Rational expectations equilibrium with conditioning on past prices: A mean-variance example, *Journal of Economic Theory* 26, 279-312.
- Ho, Thomas, and Hans Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics*, 9, 47-73.
- Ho, Thomas, and Hans Stoll, 1983, The dynamics of dealer markets under competition, *Journal of Finance* 38, 1053-1074.
- Holden, Craig, and Avanidhar Subrahmanyam, 1992, Long-lived private information and imperfect competition, *Journal of Finance* 47, 247-270.
- Hsieh, D., and A. Kleidon, 1996, Bid-ask spreads in foreign exchange markets: Implications for models of asymmetric information, in J. Frankel et al. (eds.): *The Microstructure of Foreign Exchange Markets* (University of Chicago Press, Chicago) 41-65.
- Ito, Takatoshi, Richard Lyons, and Michael Melvin, 1998, Is there private information in the FX market? The Tokyo experiment, *Journal of Finance* 53, 1111-1130.
- Kaul, A., V. Mehrotra, and R. Morck, 2000, Demand curves for stocks do slope down: New evidence from an index weights adjustment, *Journal of Finance*, 55: 893-912.
- Kim, O., and R. Verrecchia, 1991, Trading volume and price reactions to public announcements, *Journal of Accounting Research* 29, 302-321.
- Kraus, Alan, and Maxwell Smith, 1989, Market created risk, *Journal of Finance*, 44: 557-569.
- Kyle, Albert, 1989, Informed speculation with imperfect competition, *Review of Economic Studies* 56, 317-356.
- Lyons, Richard, 1995, Tests of microstructural hypotheses in the foreign exchange market, *Journal of Financial Economics*, 39: 321-351.
- Lyons, Richard, 1997, A simultaneous trade model of the foreign exchange hot potato, *Journal of International Economics* 42, 275-298.
- Lyons, Richard, 2001, *The Microstructure Approach to Exchange Rates*, MIT Press: Cambridge, Massachusetts, (chapters at [haas.berkeley.edu/~lyons](http://haas.berkeley.edu/~lyons)).
- Madhavan, Ananth, and Seymour Smidt, 1991, A Bayesian model of intraday specialist pricing, *Journal of Financial Economics* 30, 99-134.
- Madhavan, Ananth, and George Sofianos, 1997, An Empirical Analysis of NYSE Specialist Trading, *Journal of Financial Economics*, 48, 189-210.
- Madrigal, Vicente, 1996, Non-fundamental speculation, *Journal of Finance* 51, 553-578.
- Massa, M., and A. Simonov, 2003, Reputation and interdealer trading: A microstructure analysis of the treasury bond market, *Journal of Financial Markets*, 6, 99-141,

- Meese, Richard, 1990, Currency fluctuations in the post-Bretton Woods era, *Journal of Economic Perspectives*, 4: 117-134.
- Meese, Richard, and Kenneth Rogoff, 1983, Empirical exchange rate models of the seventies, *Journal of International Economics*, 14: 3-24.
- Mendelson, Haim, 1987, Consolidation, fragmentation, and market performance, *Journal of Financial and Quantitative Analysis* 22, 189-208.
- Naik, N., A. Neuberger, and S. Viswanathan, 1999, Trade disclosure regulation in markets with negotiated trades, *Review of Financial Studies*, 12: 873-900.
- Pagano, Marco, and Ailsa Roell, 1997, Front running: Market professionals as quasi-insiders, typescript, Tilburg University.
- Payne, R., 1999, Informed trade in spot foreign exchange markets: An empirical investigation, typescript, London School of Economics, *Journal of International Economics*, forthcoming.
- Radner, Roy, 1968, Competitive equilibrium under uncertainty, *Econometrica* 36, 31-58.
- Reiss, Peter, and Ingrid Werner, 1998, Does Risk Sharing Motivate Interdealer Trading? *Journal of Finance*, 53, 1657-1703.
- Roberts, H., 1967, Statistical versus classical prediction of the stock market, unpublished manuscript, Center for Research in Security Prices, University of Chicago.
- Rochet, Jean-Charles, and Jean-Luc Vila, 1994, Insider trading without normality, *Review of Economic Studies* 61, 131-152.
- Roell, A., 1990, Dual capacity trading and the quality of the market, *Journal of Financial Intermediation*, 1: 105-124.
- Saar, G., 1999, Demand Uncertainty and the Information Content of Order Flow, typescript, Johnson School, Cornell University, July.
- Scholes, Myron, 1972, The market for securities: Substitution versus price pressure and the effect of information on share price, *Journal of Business* 45, 179-211.
- Snell, Andy, and Ian Tonks, 1995, Determinants of price quote revisions on the London stock exchange, *Economic Journal* 105, 77-94.
- Spiegel, Matthew, and Avaniidhar Subrahmanyam, 1992, Informed speculation and hedging in a non-competitive securities market, *Review of Financial Studies* 5(2).
- Spiegel, Matthew, and Avaniidhar Subrahmanyam, 1995, On intraday risk premia, *Journal of Finance* 50, 319-339.
- Storesletten, Kjetil, Christopher Telmer, and Amir Yaron, 2003, Cyclical dynamics in idiosyncratic labor market risk, *Journal of Political Economy*, forthcoming.

- Subrahmanyam, Avandhar, 1991, Risk aversion, market liquidity, and price efficiency, *Review of Financial Studies* 4, 417-442.
- Vayanos, Dimitri, 2001, Strategic trading in a dynamic noisy market, *Journal of Finance*, 56, 131-171.
- Vayanos, Dimitri, 1999, Strategic trading and welfare in a dynamic market, *Review of Economic Studies* 66, 219-234.
- Viswanathan, S., and James D. Wang, 2000, Inter-dealer trading in financial markets, *Journal of Business*, forthcoming.
- Vitale, Paolo, 2000, Speculative noise trading and manipulation in the foreign exchange market, *Journal of International Money and Finance* 19, 689-712.
- Vogler, K., 1997, Risk allocation and interdealer trading, *European Economic Review* 41, 417-441.
- Vuolteenaho, T., 2002, What drives firm-level stock returns? *Journal of Finance*, 56, 233-264.
- Werner, Ingrid, 1997, A double auction model of interdealer trading, Research Paper No. 1454, Stanford University, September.
- Wurgler, Jeffrey, and Ekaterina Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks? *Journal of Business*, 75, 583-608.
- Yao, Jian, 1998, Market making in the interbank foreign exchange market, New York University Salomon Center Working Paper #S-98-4.



## Appendix A: Model Solution

This appendix repeatedly uses certain information sets and conditional expectations. To simplify notation, we present these at the outset for reference.

Information Sets:

$$\begin{aligned} \mathcal{Q}_{i1t} &\equiv \left\{ \prod_{i=1}^n \{P_{i1t}\} \right\} \\ \mathcal{Q}_{i2t} &\equiv \left\{ x_{it}, \prod_{i=1}^n \{P_{i1t}, P_{i2t}\} \right\} \\ \mathcal{Q}_{i3t} &\equiv \left\{ x_{it}, \prod_{i=1}^n \{P_{i1t}, P_{i2t}, P_{i3t}\}, T_{i2t}, T'_{i2t}, Z_t \right\} \\ \mathcal{Q}_{i4t} &\equiv \left\{ x_{it}, \prod_{i=1}^n \{P_{i1t}, P_{i2t}, P_{i3t}, P_{i4t}\}, T_{i2t}, T'_{i2t}, Z_t, T_{i3t}, T'_{i3t}, R_t \right\} \\ \mathcal{Q}_{1t} &\equiv \left\{ \prod_{i=1}^n \{P_{i1t}\} \right\} \\ \mathcal{Q}_{2t} &\equiv \left\{ \prod_{i=1}^n \{P_{i1t}, P_{i2t}\} \right\} \\ \mathcal{Q}_{3t} &\equiv \left\{ \prod_{i=1}^n \{P_{i1t}, P_{i2t}, P_{i3t}\}, Z_t \right\} \\ \mathcal{Q}_{4t} &\equiv \left\{ \prod_{i=1}^n \{P_{i1t}, P_{i2t}, P_{i3t}, P_{i4t}\}, Z_t, R_t \right\} \end{aligned}$$

The first four are the information sets available to individual dealers  $i$  at the time trades are determined in each of the four rounds. The second four are the public information sets available at the time trades are determined in each round.

### A.1. Proof of Proposition 1: Public Investors

Proposition 1 asserts that the end-of-day price is  $P_{4t} = -ah_t$ . To show this, recall that the liquidity-supplying (LS) investors have the following utility defined over intertemporal consumption  $c_t$  (per equation 1):

$$(1) \quad U_t = \sum_{\tau=t}^{\infty} -\delta^{\tau-t} \exp(-\gamma c_{\tau})$$

We begin by conjecturing a value function, which we show below is consistent with optimizing behavior on the part of LS investors:

$$V_t = -\alpha \exp(-\hat{\gamma} W_t - \psi h_t^2)$$

where  $W_t$  is the LS investors' nominal wealth at the end of day  $t$  and  $h_t$  is the total holding of the risky asset at the end of day  $t$  (defined in equation 9). We need to determine the conditions under which  $h_t$  is willingly held by the LS investors.

Given the proposed price function  $P_{4t} = -ah_t$  in text proposition 1, our task is to begin with the Bellman equation corresponding to the maximum of equation (1) and derive explicit

expressions for the three coefficient values  $\hat{\gamma}$ ,  $\psi$ , and  $\alpha$ , in this conjectured value function, as well as an expression for the parameter  $a$  in the price function. We shall show that we must have:

$$\hat{\gamma} = \left( \frac{r}{1+r} \right) \gamma$$

$$\psi = \left( \frac{1}{2n\sigma_x^2} \right) + (1+r)a\hat{\gamma} - \left( \frac{(a\hat{\gamma} + \frac{1}{n\sigma_x^2})[(1+r)a - \sigma_R^2\hat{\gamma}]}{2a} \right) - \left( \frac{\sigma_R^2\hat{\gamma}}{2} \right)$$

$$\alpha = \left( \frac{\hat{\gamma}}{\gamma} \alpha \sqrt{1+2n\sigma_x^2\psi} \right)^{\frac{1}{1+r}} + \left( \frac{\hat{\gamma}}{\gamma} \alpha \sqrt{1+2n\sigma_x^2\psi} \right)^{-\frac{r}{1+r}} \left( \frac{\alpha}{1+r} \sqrt{1+2n\sigma_x^2\psi} \right)$$

and

$$a^2\hat{\gamma} - ma \left( \frac{r}{n\sigma_x^2} + 2\psi + 2r\psi \right) + \left( \frac{1}{n\sigma_x^2} + 2\psi \right) \hat{\gamma} \sigma_R^2 = 0$$

To prove these conditions, write down the Bellman equation:

$$V_t = \text{Max}_{\{c_t, D_t\}} -\exp(-\gamma c_t) - \delta E_t \left[ \alpha \exp(-\hat{\gamma} W_{t+1} - \psi h_{t+1}^2) \right]$$

where

$$W_{t+1} = (1+r)(W_t - c_t) + D_t (P_{t+1} + R_{t+1} - (1+r)P_t)$$

and where we have used  $D_t$  to denote the LS investors' demand for the risky asset. The first order condition with respect to  $c_t$  is:

$$\gamma \exp(-\gamma c_t) = \delta \hat{\gamma} (1+r) E_t \left[ \alpha \exp(-\hat{\gamma} W_{t+1} - \psi h_{t+1}^2) \right]$$

Notice that the consumption decision is unaffected by the investment decision  $D_t$ , due to CARA utility. To get an explicit expression for the right-hand side, we calculate the following expectation with respect to the two random variables  $R_{t+1}$ , and  $x_{t+1}$ , both of which are normally distributed with mean zero and respective variances  $\sigma_R^2$  and  $\sigma_x^2$ :

$$(A1) \quad E_t \left[ -\exp(-\hat{\gamma} W_{t+1} - \psi h_{t+1}^2) \right] =$$

$$- \left( \sqrt{1+2n\sigma_x^2\psi} \right) \exp \left( -\hat{\gamma} (1+r)(W_t - c_t) + \frac{\left( a\hat{\gamma} D_t + \frac{h_t}{n\sigma_x^2} \right)^2}{4\psi + 2/(n\sigma_x^2)} + \left( \frac{\hat{\gamma}^2 D_t^2 \sigma_R^2}{2} \right) - (1+r)a\hat{\gamma} D_t h_t - \left( \frac{h_t^2}{2n\sigma_x^2} \right) \right)$$

Maximizing with respect to the choice of risky asset demand  $D_t$ , we get

$$D_t = \left( \frac{a(1+r) - \left(\frac{a}{n\sigma_x^2}\right) \left(\frac{1}{n\sigma_x^2} + 2\psi\right)^{-1}}{\hat{\gamma}\sigma_R^2 + a^2\hat{\gamma} \left(\frac{1}{n\sigma_x^2} + 2\psi\right)^{-1}} \right) h_t$$

For market clearing, we must have  $mD_t = h_t$ , so:

$$1 = m \frac{a(1+r) - \left(\frac{a}{n\sigma_x^2}\right) \left(\frac{1}{n\sigma_x^2} + 2\psi\right)^{-1}}{\hat{\gamma}\sigma_R^2 + a^2\hat{\gamma} \left(\frac{1}{n\sigma_x^2} + 2\psi\right)^{-1}}$$

This is equivalent to the expression for the pricing parameter “ $a$ ” presented above. To get the coefficient  $\psi$  on  $h_t^2$  in the value function, we collect terms in equation (A1) involving  $h_t^2$ :

$$\psi = \frac{1}{2n\sigma_x^2} + \frac{a(1+r)\hat{\gamma}}{m} - \frac{\left(\frac{a\hat{\gamma}}{m} + \frac{1}{n\sigma_x^2}\right) \left(a(1+r) - \frac{\sigma_R^2\hat{\gamma}}{m}\right)}{2a} - \frac{\sigma_R^2\hat{\gamma}^2}{2m^2}$$

Substituting the expected value function in the next period back into the Bellman equation, we get the expression for  $\alpha$ :

$$\alpha = \left( \frac{\hat{\gamma}}{\gamma} \alpha \sqrt{1 + 2n\sigma_x^2\psi} \right)^{\frac{1}{1+r}} + \left( \frac{\hat{\gamma}}{\gamma} \alpha \sqrt{1 + 2n\sigma_x^2\psi} \right)^{\frac{r}{1+r}} \left( \frac{\alpha}{1+r} \right) \sqrt{1 + 2n\sigma_x^2\psi}$$

It is easy to show that when the coefficient of risk aversion  $\gamma$  is sufficiently small, there exists a positive solution for the parameters. Finally, that a value function with this simple exponential form exists ensures that the linear equilibrium pricing rule described in proposition 1 also exists (recall that the mean cash flow on the risky asset is zero). Q.E.D.

## A.2. Proof of Propositions 2 and 3: Price Determination in the First Three Rounds

First, consider why in any given trading round all dealers  $i \in \{1, \dots, n\}$  choose to offer the same price. As shown in a similar setting in Lyons (1997), this is necessary to avoid unbounded arbitrage demands given that quotes are singleton prices and available to all participants (see quoting rules R1-R3 and trading rules R4-R6). For risk-averse dealers, an equilibrium that includes such arbitrage trading cannot be optimal. A common price, then, must depend only on commonly observed information, i.e., public information.

The equilibrium level of the round-one price is  $P_{1t} = P_{1t} = E[P_{2t} | \Omega_{1t}]$ , i.e., quotes to customers in the first round are an unbiased estimate of the round-two price (conditional on public information available at the time of quoting  $\Omega_{1t}$ ). Two main features of the model produce this result: dealers cannot forecast customers' trades and dealers are risk averse. To prove this formally, based on the value function from the previous appendix section (with positive constants  $A_1$  and  $A_2$ ) one can demonstrate that:

$$\begin{aligned}
 \text{(A1)} \quad E[U_i(P_{1t})] &= E[E[U_i | \Omega_{1t}]] = E[-A_1 \exp(-\hat{\gamma}(P_{1t} - P_{2t})x_{it} - A_2x_{it}^2)] \\
 &= E[E[-A_1 \exp(-\hat{\gamma}(P_{1t} - P_{2t})x_{it} - A_2x_{it}^2) | x_{it}^2]] \\
 &\leq E[U_i(E[P_{2t} | \Omega_{1t}])] \\
 &= E[-A_1 \exp(-A_2x_{it}^2)] \\
 &= E[-A_1 \exp(-\hat{\gamma}(P_{2t} - P_{2t})E[x_{it} | x_{it}^2] - A_2x_{it}^2)].
 \end{aligned}$$

Here, the inequality in the third line arises because absorbing a random position of  $-x_{it}$  at a price  $P_{1t}$  that is conditionally biased relative to the unwinding price  $P_{2t}$  can only add to risk (with no effect on expected return). The inequality is strict for any price  $P_{1t} \neq E[P_{2t} | \Omega_{1t}]$ , so it is optimal for the dealer not to deviate from the proposed equilibrium strategy.

The equilibrium level of the round-two price is  $P_{2t} = P_{2t} = E[P_{3t} | \Omega_{2t}]$ , i.e., quotes to other dealers in the second round are an unbiased estimate of the round-three price. To see this, note that at the beginning of the second round of trading—the first purely interdealer round—the expected current holding of dealers conditional on public information is still zero (since there is no new public information useful for estimating dealer positions). Since the round-two market must clear among dealers alone, the only price not expected to generate excess dealer demand is the expected future price. Specifically, the price  $P_{2t}$  in proposition 2 is pinned down by the relation:

$$(A2) \quad \sum_{i=1}^n (E[x_{it} | \Omega_{2t}] + E[D_{i2t}(P_{2t}) | \Omega_{2t}]) = 0$$

where  $\Omega_{2t}$  is public information available for quoting. (Since  $P_{2t}$  is common across dealers  $i$  it is necessarily conditioned on public information only.) At the time of quoting in round two there is nothing in  $\Omega_{2t}$  that helps estimate  $x_{it}$  so  $E[x_{it} | \Omega_{2t}] = 0$ . The only value of  $P_{2t}$  for which  $E[D_{i2t}(P_{2t}) | \Omega_{2t}] = 0$  is  $P_{2t} = E[P_{3t} | \Omega_{2t}]$  since  $D_{i2t}(E[P_{3t} | \Omega_{2t}]) = 0$  and  $D'_{i2t} < 0$ .

As proposition 3 asserts, in the third round a bias in  $P_{3t}$  is necessary for the analogue of Eq. (A2) to hold:

$$(A3) \quad \sum_{i=1}^n (E[x_{it} | \Omega_{3t}] + E[D_{i3t}(P_{3t}) | \Omega_{3t}]) = 0$$

First, consider the term  $E[x_{it} | \Omega_{3t}]$ . Given  $P_{1t}$ ,  $P_{2t}$  and  $P_{3t}$  are common across dealers and conditioned only on public information, the only variable in  $\Omega_{3t}$  relevant for determining  $P_{3t}$  is  $Z_t$ , interdealer order flow from round two. We have:

$$E[x_{it} | \Omega_{3t}] = E[x_{it} | Z_t] = E[x_{it} | \sum_i \beta_2 x_{it}] = (n\beta_2)^{-1} Z_t$$

To determine  $E[D_{i3t}(P_{3t}) | \Omega_{3t}]$  in Eq. (A3), we use normality and exponential utility to write:

$$E[D_{i3t} | \Omega_{3t}] = E[P_{4t} + R_t - (1+r)P_{3t} | \Omega_{3t}] / (\hat{\gamma}\sigma_R^2) = -(1+r)\lambda Z_t / (\hat{\gamma}\sigma_R^2).$$

These expressions in Eq. (A3) imply:

$$\lambda = \hat{\gamma}\sigma_R^2 / [(1+r)n\beta_2]$$

with  $\lambda > 0$  unambiguously.

### **A.3. Proof of Proposition 4: Optimal Dealer Trading Strategies**

The derivation of dealer trading strategies has 3 steps. First, we determine the round-three trading strategy given actions in earlier rounds. (We do not start with round-four trading because dealers' round-four trading is passive: they simply take the other side of LS-customer trades, ending the day with a net position of zero.) Next, we determine the expected utilities of possible trading strategies in round 2. Finally, we solve this maximization problem.

#### **A.3.1. Step One: Maximization in the Third Round**

To solve the dynamic programming problem, we first determine the dealer's round-three

desired position for use in the round-two first-order condition. Under normality and negative exponential utility, the round-three desired position takes the well-known form:

$$D_{i3t} = E[P_{4t} + R_t - (1+r)P_{3t}] / (\hat{\gamma}\sigma_R^2) = E[P_{4t} - (1+r)P_{3t}] / (\hat{\gamma}\sigma_R^2)$$

where  $\hat{\gamma}$  is the intertemporal coefficient of absolute risk aversion and  $\sigma_R^2$  denotes  $\text{Var}[R_t]$ .

Notice from proposition 3 that we have:

$$P_{3t} = E\left[\frac{P_{4t}}{1+r} \mid \Omega_{3t}\right] + \lambda Z_t = \frac{-a(h_{t-1} - Z_t / \beta_2)}{1+r} + \lambda Z_t$$

$$P_{4t} = -a[h_{t-1} - x_{it} - (Z_t - T_{i2t}) / \beta_2]$$

Thus, we have:

$$P_{4t} - (1+r)P_{3t} = a(x_{it} - T_{i2t} / \beta_2) - (1+r)\lambda Z_t$$

and

$$(A4) \quad D_{i3t} = [a(x_{it} - T_{i2t} / \beta_2) - (1+r)\lambda Z_t] / (\hat{\gamma}\sigma_R^2)$$

Omitting terms unrelated to  $D_{i2t}$ , we can write the dealers' problem as:

$$(A5) \quad \underset{D_{i2t}}{\text{Max}} \quad E\left[-\exp\left(-\hat{\gamma}(1+r)(D_{i2t} - T'_{i2t})(P_{3t} - P_{2t}) - (P_{4t} - (1+r)P_{3t})\sigma_R^{-2}(P_{4t} + R_t - (1+r)P_{3t})\right) \mid \Omega_{i2t}\right]$$

### A.3.2. Step Two: Expected Dealer Utility after Second Round Trading

We can use the moment generating function for the normally distributed variable  $R_t$  to re-express the last term in the exponential function so that the full problem is now:

$$(A6) \quad \underset{D_{i2t}}{\text{Max}} \quad E\left[-\exp\left(-\hat{\gamma}(1+r)(D_{i2t} - T'_{i2t})(P_{3t} - P_{2t}) - \frac{(P_{4t} - (1+r)P_{3t})^2}{2\sigma_R^2}\right) \mid \Omega_{i2t}\right]$$

which leaves the objective function focused on the random variable  $P_{3t}$ .

Now, notice that interdealer order flow  $Z_t$  can be written as the sum of two components, dealer  $i$ 's interdealer order plus the sum of all the other dealers' interdealer orders:

$$Z_t = T_{i2t} + (n-1)T'_{i2t}$$

where

$$T'_{i2t} = \frac{\sum_{i \neq i} \beta_2 x_{i't}}{n-1} = \frac{Z_t - T_{i2t}}{n-1}$$

and the product  $(n-1)T'_{i2t}$  arises because dealer  $i$  receives a share  $1/(n-1)$  of all other dealers' outgoing trades (by trading rule R6). Moreover, we have:

$$P_{3t} = \frac{-a(h_{t-1} - Z_t / \beta_2)}{1+r} + \lambda Z_t,$$

$$P_{2t} = E[P_{3t} | \Omega_{2t}] = \frac{-ah_{t-1}}{1+r}.$$

As a result, the price change is:

$$P_{3t} - P_{2t} = \left[ \frac{a}{(1+r)\beta_2} + \lambda \right] Z_t.$$

After substituting the expressions for  $P_{2t}$ ,  $P_{3t}$ , and  $P_{4t}$  into (A6) and taking expectations with respect to  $Z_t$ , dealer  $i$ 's expected utility as a function of his speculative demand  $D_{i2t}$  is:

$$\int_{-\infty}^{\infty} \left[ -\exp \left( -\hat{\gamma}(1+r) \left( D_{i2t} - \frac{Z_t - T_{i2t}}{n-1} \right) \left( \frac{a}{(1+r)\beta_2} + \lambda \right) Z_t - \left( \frac{[a(x_{it} - \frac{T_{i2t}}{\beta_2}) - (1+r)\lambda Z_t]^2}{2\sigma_R^2} \right) - \left( \frac{[Z_t - T_{i2t}]^2}{2(n-1)\beta_2^2\sigma_x^2} \right) \right) \right] dZ_t$$

### A.3.3. Step Three: Maximization in the Second Round

Notice that  $T_{i2t} = \beta_2 x_{it} = D_{i2t} + x_{it}$ . After ignoring irrelevant terms, with some algebra the maximization problem reduces to:

$$\text{Max}_{D_{i2t}} \frac{B^2 - 4AC}{4A}$$

where  $A$ ,  $B$ , and  $C$  are the second, first, and zero-order coefficients of  $Z_t$ , respectively:

$$A = \frac{\hat{\gamma}}{n-1} \left( \frac{a}{\beta_2} + \lambda(1+r) \right) - \left( \frac{[(1+r)\lambda]^2}{2\sigma_R^2} \right) - \left( \frac{1}{2(n-1)\beta_2^2\sigma_x^2} \right)$$

$$B = -\hat{\gamma} \left( \frac{a}{\beta_2} + \lambda(1+r) \right) \left( D_{2it} + \frac{D_{2it} + x_{it}}{n-1} \right) + \left( \frac{D_{2it} + x_{it}}{(n-1)\beta_2^2\sigma_x^2} \right) + \left( \frac{a}{\sigma_R^2} (1+r)\lambda \right) \left( x_{it} - \frac{D_{2it} + x_{it}}{\beta_2} \right)$$

$$C = -\frac{a^2}{2\sigma_R^2} \left( x_{it} - \frac{D_{2it} + x_{it}}{\beta_2} \right)^2 - \left( \frac{(D_{2it} + x_{it})^2}{2(n-1)\beta_2^2\sigma_x^2} \right).$$

The maximization problem reduces to a quadratic maximization problem. Assuming that  $A$  is negative (which we return to below), we take the first-order condition for the maximization problem (noting that  $A$  does not depend on  $D_{i2t}$ ) to get:

$$(A7) \quad 2B \frac{\partial B}{\partial D_{i2t}} - 4A \frac{\partial C}{\partial D_{i2t}} = 0.$$

Noting that:

$$D_{i2t} = (\beta_2 - 1)x_{it}$$

and

$$\lambda = \frac{\hat{\gamma}\sigma_R^2}{(1+r)n\beta_2}$$

substituting these expressions for  $\lambda$  and  $D_{i2t}$  in (A7) gives the following quadratic equation:

$$(A8) \quad \left(-\hat{\gamma}\left[a + \frac{\hat{\gamma}\sigma_R^2}{n}\right][n\beta_2 - n + 1] + \frac{1}{\sigma_x^2}\right) \left(\hat{\gamma}\left[a + \frac{\hat{\gamma}\sigma_R^2}{n}\right]n\beta_2 + \frac{1}{\sigma_x^2} - \frac{(n-1)a}{\sigma_R^2} \left[\frac{\hat{\gamma}\sigma_R^2}{n}\right]\right) - \left(\frac{2A(n-1)\beta_2^2}{\sigma_x^2}\right) = 0$$

When the number of dealers  $n$  goes to infinity and the pricing coefficient  $a$  goes to zero, the quadratic equation (A8) reduces to:

$$\hat{\gamma}^2\sigma_R^2(\beta_2 - 1) \left(\hat{\gamma}^2\sigma_R^2\beta_2 + \frac{1}{\sigma_x^2}\right) - \frac{1}{\sigma_x^4} = 0.$$

There exists a solution  $\beta_2$  in the interval  $(1, \infty)$  when  $n$  is infinity and  $a$  is zero. Since the coefficients of  $\beta_2$  in (A8) is continuous in  $n$  and  $a$ , there exists a unique solution for  $\beta_2$  in the interval  $(1, \infty)$ , for large  $n$ . Further, it is easy to check that the assumption that  $A$  is negative is satisfied for large  $n$ . Q.E.D.



## Appendix B: Estimation Details

The model to be estimated takes the form:

$$\begin{aligned}
 \text{(B1)} \quad & P_t = P_t^{PI} + P_t^{II} + w_t \\
 \text{(B2)} \quad & \Delta P_t^{PI} = \varepsilon_t \\
 \text{(B3)} \quad & \Delta P_t^{II} = \Gamma(L)e_t \\
 \text{(B4)} \quad & Z_t = \Theta(L)e_t
 \end{aligned}$$

where  $\Gamma(L)$  and  $\Theta(L)$  are polynomials in the lag operator and  $w_t$ ,  $\varepsilon_t$ , and  $e_t$  are mean zero shocks.  $P_t$  is the last DM purchase price for dollars during observation interval  $t$  and  $Z_t$  is interdealer order flow over the same interval.

In our theoretical model, LD customer orders never affect prices before being reflected (and publicly observed) in interdealer order flow. This would imply that all price implications of aggregate LD customer flows can be identified from the effects of current and lagged interdealer flow (see Proposition 3). In actual markets, the timing of customer-flow effects on prices is a bit subtler. In particular, some dealers (i.e. those working for large money center banks) observe the trading of a large customer base that enables inferences about other banks' LD customer flows *before* that information is reflected in interdealer flow. Because quotes in actual FX markets are not necessarily observable to all participants (in contrast to our model), there is some room for these dealers to revise their quotes in anticipation of interdealer flow. (See Evans 2002 for a further discussion of quote transparency in the interdealer market.)

To make room in our empirical model for this possibility, we model shocks to measured interdealer flow over interval  $t$ ,  $e_t$ , as follows. Let the shock to order flow during interval  $t$  have two components:  $e_t = e_{1,t} + e_{2,t}$ . We assume that  $e_{1,t}$  shocks contain information on LD customer flows that is embedded in prices one period before being manifested in interdealer flow. (With a period length of 5 minutes, this “early transmission” channel is thus less than half of the typical inventory half-life of about 10 minutes; see Lyons, 1995.) LD flow information that becomes embedded in prices after being manifested in interdealer flow is represented by  $e_{2,t}$ . We can now express the dynamics of measured order flow and II price changes by:

$$(B5) \quad \Delta P_t^{II} = \Gamma(L)(e_{1,t+1} + e_{2,t})$$

$$(B6) \quad Z_{t+1} = \Theta(L)(e_{1,t+1} + e_{2,t+1})$$

For estimation, we assume that  $\Gamma(L) = (b_0 + b_1L + b_2L^2)/(1 - a_1L - a_2L^2)$  and  $\Theta(L) = (1 + d_1L + d_2L^2)/(1 - c_1L - c_2L^2)$ , so the model in (B1), (B2), (B5), and (B6) can be written in state-space form as:

$$(B7) \quad \zeta_t = A\zeta_t + B\eta_t$$

$$\xi_t = C\zeta_t$$

with  $E[\eta_t] = 0$  and  $E[\eta_t \eta_t'] = \Omega$ , where:

$$\zeta_t' = [\Delta P_t - \Delta P_t^{II} \quad \Delta P_t^{II} \quad \Delta P_{t-1}^{II} \quad Z_{t+1} \quad Z_t \quad e_{1,t+1} \quad e_{2,t+1} \quad e_{1,t} \quad e_{2,t} \quad e_{2,t-1} \quad w_t]$$

$$\xi_t' = [\Delta P_t \quad Z_{t+1}]$$

$$\eta_t' = [e_{1,t+1} \quad e_{2,t+1} \quad \varepsilon_t \quad w_t]$$

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & a_1 & a_2 & 0 & 0 & b_1 & b_0 & b_2 & b_1 & b_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_1 & c_2 & d_1 & d_1 & d_2 & d_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 & 1 \\ b_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \Omega = \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & 0 & \sigma_w^2 \end{bmatrix}.$$

The GMM estimates of the parameters ( $a_i$ ,  $b_i$ ,  $c_i$ ,  $d_i$ , and  $\sigma_i$ ) are computed by equating the sample moments,  $Cov[\xi_t, \xi'_{t-j}]$  for  $j=0,1,\dots,12$ , against their counterparts implied by the state-space form in equation (B7):  $C\Psi(j)C'$  where  $\Psi(j)=A\Psi(j-1)$  with  $\Psi(0)=vec^{-1}[(I-A\otimes A)^{-1}vec(B\Omega B')]$ . (Evans 2002 uses a similar estimation procedure and further details are available there.) To conserve space, Table 1 reports estimates of specifications in which  $b_2=0$ : estimates of the model without this restriction were very similar to those reported, and the estimates of  $b_2$  were statistically insignificant.

To compute the contribution of inventory information to the variance of II price changes, we first need to decompose II price changes into their inventory and non-inventory components:

$$\begin{aligned} \Delta P_t^{II} &= \Gamma(L)e_t \\ &= (\Gamma(L) - \Gamma(1))e_t + \Gamma(1)(1 - L^\tau)e_t + \Gamma(1)e_{t-\tau} \\ &= (\Gamma(L) - \Gamma(1)L^\tau)e_t + \Gamma(1)e_{t-\tau} \\ &= \bar{\Gamma}(L)e_t + \Gamma(1)e_{t-\tau} \end{aligned}$$

where  $e_t = e_{1,t} + e_{2,t}$ . We identify the transitory component of the  $\Delta P_t^{II}$  process as  $\Delta P_t^{trans} = \bar{\Gamma}(L)e_t$ . This decomposition is slightly more general than the one given in the text because it allows a lag of  $\tau$  periods before the permanent effects of the shock  $e_t$  are reflected in prices. Table 1 reports results for the  $\tau=3$  case (i.e., a lag of 15 minutes), but the results we obtain for the cases where  $\tau=0,1$ , and 2 are very similar. In particular, the estimates of  $\bar{\Gamma}(0)$  are statistically significant in every case.

To compute the variance contributions, we use the GMM estimates to write the  $\Delta P_t^{II}$  and  $\Delta P_t^{trans}$  processes as a new state space system:

$$\begin{aligned} \mathcal{Z}_t &= \mathcal{A}\mathcal{Z}_{t-1} + \mathcal{B}e_t \\ \mathcal{Y}_t &= \mathcal{C}\mathcal{Z}_t \end{aligned}$$

where  $\mathcal{Y}'_t = [\Delta P_t^H \ \Delta P_t^{trans}]$ . With this system, we can now compute  $Var[\Delta^k P_t^H] = Var[h_1 \sum_{i=0}^{k-1} \mathcal{Y}'_{t-i}]$  and  $Var[\Delta^k P_t^{trans}] = Var[h_2 \sum_{i=0}^{k-1} \mathcal{Y}'_{t-i}]$ , where  $\Delta P_t^H = h_1 \mathcal{Y}'_t$  and  $\Delta P_t^{trans} = h_2 \mathcal{Y}'_t$ . For the  $\tau=3$  case, the state-space form is given by:

$$\mathcal{X}'_t = [\Delta P_t^H \ \Delta P_{t-1}^H \ e_t \ e_{t-1} \ e_{t-2} \ e_{t-3}],$$

$$\mathcal{A} = \begin{bmatrix} a_1 & a_2 & b_1 & b_2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} b_0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathcal{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -\Gamma(1) \end{bmatrix}.$$